

Univerzita Karlova v Praze

Přírodovědecká fakulta

Studijní program: Biochemie



Acidobazické a tautomerní rovnováhy bází nukleových kyselin

Acid-base equilibria and tautomerism of nucleobases

Bakalářská práce

Jakub Kovács

Vedoucí práce: RNDr. Helena Dračínská, Ph.D.

Konzultant: RNDr. Martin Dračínský, Ph.D.

Praha 2016

Prohlašuji, že jsem tuto bakalářskou práci vypracoval samostatně pod vedením RNDr. Martina Dračínského, Ph.D. a RNDr. Heleny Dračínské, Ph.D. a že jsem všechny použité prameny řádně citoval.

V Praze, dne:

Podpis:

Rád bych touto cestou poděkoval svému konzultantovi RNDr. Martinovi Dračínskému, Ph.D. za trpělivost a ochotu během vedení této práce a velmi cenné rady při jejím vypracování.

Také bych velmi rád poděkoval RNDr. Heleně Dračínské, Ph.D. za cenné rady při sepisování práce, Radkovi Pohlovi za trpělivost a čas při zaškolení do experimentálních metod a celému vědecko-servisnímu týmu NMR spektroskopie za vytvoření velmi příjemného pracovního prostředí.

Abstrakt

Nukleové kyseliny a jejich složky patří mezi nejdůležitější molekuly pro život. Přesunem vodíkového atomu po skeletu pyrimidinových a purinových bází nukleových kyselin vznikají různé tautomerní formy těchto bází. V závislosti na pH prostředí mohou být navíc báze nukleových kyselin protonovány nebo deprotonovány. Protonace/deprotonace bází nukleových kyselin se účastní poměrně velkého množství biochemických dějů. Jedná se hlavně o procesy spojené s regulací, replikací či genovou expresí, ve kterých hrají ionizované báze důležitou úlohu ve stabilizaci sekundárních struktur nukleových kyselin, či v acidobazické katalýze autolytických reakcí jejich fosfodiesterové kostry. Sekundární struktury stabilizované ionizovanými bázemi jsou navíc spojené s různými genetickými chorobami. Protonace a deprotonace jsou však velmi rychlé a dynamické procesy, které je velmi náročné experimentálně detekovat. Tato práce se proto zabývá vypracováním spolehlivé metodiky, pomocí které by bylo možné poměrně rychle zjistit, na kterém atomu dusíku pyrimidinových bází dochází k protonaci. V experimentech byla využita NMR spektroskopie, pomocí které byly nejprve zkoumány pyrimidinové báze, u kterých je místo protonace jednoznačně určeno strukturou. Byly získány experimentální NMR parametry pro neutrální i ionizované formy. NMR parametry pro neutrální a ionizované formy těchto látek byly následně vypočteny pomocí kvantově chemických výpočtů. Způsob výpočtu byl optimalizovaný tak, aby výsledky co nejlépe korelovaly s experimentálními hodnotami. Metoda korelace výpočet – experiment byla otestována při detekci místa protonace bází, u kterých toto místo není jasně dáno strukturou. Na základě výsledků byla navržena optimální metoda pro kvantově chemický výpočet NMR parametrů.

Klíčová slova: pyrimidin, báze nukleových kyselin, NMR spektroskopie, ionizace, DFT výpočty

Abstract

Nucleic acids and their components belong to the most important molecules essential for life. Hydrogen atom migration across the skeleton of purine or pyrimidine bases of nucleic acids gives rise to their various tautomeric forms. Furthermore, depending on pH of the medium, these bases might become protonated or deprotonated. Protonation/deprotonation of nucleic acid bases is involved in many biochemical processes. These include mainly regulation of replication and gene expression by means of specific secondary structure stabilization or acid-base catalysis of phosphodiester backbone cleavage. Additionally, secondary structures that include ionized bases are also associated with various genetic diseases. However, protonation and deprotonation are very fast dynamic processes and their experimental detection is difficult. Therefore, this work focuses on development of reliable method for relatively fast detection of pyrimidine base derivatives primary protonation site. At first, pyrimidine bases with protonation site determined by structure were analyzed in both neutral and ionized state using NMR spectroscopy. In the next step, NMR parameters for both neutral and ionized forms were calculated using quantum-chemical computational methods. Computational method was optimized to produce results which correlate with experimental values. Method based on correlation of optimized calculation with NMR experiment was used to detect primary protonation site of pyrimidine bases with more than one possible site. Based on the results, an optimal method for quantum chemical computation of NMR parameters was designed.

(In Czech)

Keywords: pyrimidine, nucleic acid base, NMR spectroscopy, ionization, DFT computation

Obsah

Seznam použitých zkratk	7
1 Úvod	8
1.1 Nukleové kyseliny	8
1.1.1 Strukturní jednotky	8
1.1.2 Struktura a funkce	9
1.2 Modifikace složek nukleových kyselin	11
1.2.1 Modifikace chemické	11
1.2.2 Modifikace fyzikálně-chemické	12
1.2.3 Protonace a deprotonace	13
1.3 Vlivy ionizace bází na strukturu nukleových kyselin	14
1.3.1 pH prostředí jako evoluční faktor	14
1.3.2 Hoogsteenův model	15
1.3.3 Triplexy stabilizované Hoogsteenovým párováním	16
1.3.4 Další struktury ionizovaných bází	18
1.4 Ionizované a minoritní tautomerní formy v biologických procesech	18
1.4.1 H-DNA a regulace genové exprese	18
1.4.2 H-DNA a regulace replikace DNA	20
1.4.3 RNA přepínače	21
1.4.4 RNA katalýza	22
1.4.5 Poruchy spojené s ionizovanými formami	25
1.4.6 Potenciální terapeutické a biologické využití	27
2 Cíle práce	29
3 Materiál a metody	30
3.1 Materiál	30
3.2 Přístroje	30
3.3 Software	30
3.4 Metody	31
3.4.1 Příprava roztoků bází	31
3.4.2 NMR spektroskopie	32
3.4.3 Optimalizace geometrie a výpočet NMR parametrů	32
4 Výsledky a diskuze	34

4.1 Přiřazení signálů pomocí 1D a 2D NMR spekter	34
4.1.1 ^1H NMR	34
4.1.2 ^{13}C NMR	36
4.1.3 ^1H , ^{13}C HMBC NMR	37
4.1.4 ^1H , ^{15}N HMBC NMR	38
4.2 Titrace zásadou/kyselinou a změny chemických posunů modelových molekul	39
4.3 Vývoj metody detekce místa protonace korelací výpočet – experiment	43
4.3.1 Výběr vhodné výpočetní metody	43
4.3.2 korelace výpočet – experiment pro ^{13}C NMR parametry neutrální molekuly	46
4.3.3 Kontrola odchylky vybrané metody na látce se strukturně definovaným místem protonace	46
4.4 Testování metody	48
4.4.1 Detekce místa protonace 2-amino-4-methylpyrimidinu porovnáním experimentu s výpočtem	48
4.4.2 Detekce místa protonace báze s omezeným množstvím experimentálně dostupných dat	52
5 Shrnutí	57
6 Seznam použité literatury	58

Seznam použitých zkratek

ATP	adenosintrifosfát
APT	z angl. attached proton test
DFT	z angl. density functional theory
DNA	deoxyribonukleová kyselina (z angl. deoxyribonucleic acid)
Glms	ribozym netranslatované oblasti glukosamin-6-fosfát synthetasy
GlcN6P	glukosamin-6-fosfát
HDV	z angl. Hepatitis delta virus
HIV	z angl. Human imunodeficiency virus
HMBC	z angl. heteronuclear multiple bond correlation
mRNA	messengerová RNA (z angl. messenger RNA)
NAD	nikotinamidadenindinukleotid
NMR	nukleární magnetická rezonance (z angl. nuclear magnetic resonance)
OxyTPP	4'-keto-N1H-oxythiamín pyrofosfát
RNA	ribonukleová kyselina (z angl. „ribonucleic acid“)
SV 40	z angl. „Simian vacuolating virus 40“
TFO	triplex formující oligonukleotidy (z angl. triplex forming olegonucleotides)
TPP	thiaminpyrofosfát
tRNA	transferová RNA (z angl. „transfer RNA“)
UOCHB AVČR	Ústav organické chemie a biochemie Akademie věd České Republiky
WC	Watson-Crick

1 Úvod

1.1 Nukleové kyseliny

Roku 1865 Johann Gregor Mendel formuloval zákony dědičnosti. Čtyři roky po něm švýcarský lékař Friedrich Meischer izoloval z jader bílých krvinek obsažených v hnisu látky bohaté na fosfor, které nazval nuklein. Brzy se však ukázalo, že tyto látky jsou kyselé povahy a byly přejmenovány na nukleové kyseliny. Až o několik desítek let později, s objevem Jamese Watsona, Francise Cricka, Mauritia Wilkinse a Rosalindy Franklinové, byla odhalena jejich struktura a biologická funkce. Nukleové kyseliny jsou biopolymery, které uchovávají genetickou informaci organismu, a proto je lze považovat za jedny z nejdůležitějších biologických makromolekul. Podle strukturních jednotek, které je formují, je dělíme do dvou základních skupin: DNA a RNA.

1.1.1 Strukturní jednotky

Základní strukturní jednotkou nukleových kyselin jsou nukleotidy. Každá taková monomerní jednotka obsahuje pětiuhlíkatý monosacharid: D-ribosu, v případě RNA, a D-2-deoxyribosu, v případě DNA. Další složku tvoří zbytky kyseliny fosforečné. Fosfátové skupiny, které propojují uhlíky 3 a 5 sousedních pentos fosfodiesterovou vazbou, tvoří pentosa-fosfátovou páteř nukleových kyselin u obou typů kyselin stejným způsobem. Poslední, a z hlediska uchování genetické informace nejdůležitější, složkou jsou nukleové báze. Z chemického hlediska se jedná o deriváty dvou dusíkatých heterocyklů – purinu a pyrimidinu. Mezi purinové deriváty patří adenin (A) a guanin (G), pyrimidinové deriváty jsou cytosin (C), thymin (T) a uracil (U). Thymin najdeme výhradně v DNA, zatímco jeho o methyl ochuzená forma uracil se nachází jen v RNA. Tyto báze však mohou být různě modifikovány. Proto v molekulách nukleových kyselin najdeme velké množství jejich derivátů. Nukleotidy mají i jiné důležité funkce, mezi které patří například interakce s enzymy a dalšími proteiny jako komponenty kofaktorů a signálních molekul. Adenosintrifosfát (ATP) je jedna z nejdůležitějších makroergních sloučenin zodpovědná za zásobování buněk energií. Jsou také strukturními složkami kofaktorů nezbytných pro buněčný metabolismus, jako například nikotinamidadenindinukleotid (NAD).

1.1.2 Struktura a funkce

DNA

Podle původního Watson-Crickovského modelu jsou jednotlivá vlákna uspořádána do šroubovicové struktury stabilizované vodíkovými můstky mezi bázemi jednotlivých řetězců, a to právě tak, že se páruje pyrimidinová báze jednoho řetězce s purinovou bází řetězce druhého. Vždy se však jedná o páry strukturně komplementární. Watson a Crick předpokládali, že se adenin páruje s thyminem a cytosin s guaninem. V obou případech ale předpokládali tvorbu 2 vodíkových můstků. Krátce po jejich objevu, navrhl Pauling existenci tří můstků mezi cytosinem a guaninem, což bylo později potvrzeno.[1] Specifická komplementarita párování je klíčová pro uchování genetické informace, protože umožňuje replikaci struktur nukleových kyselin v přesně stejném pořadí jednotlivých bází. Vzniká tak pravotočivá dvoušroubovice složená ze dvou antiparalelních vláken. Nejčastější formou je B-DNA s deseti páry bází na jednu otáčku.[2] Tato forma však není jediná. Jí velmi podobnou formou je A-DNA, pravotočivá struktura také stabilizovaná WC páry bází. Její odlišnost spočívá jen v malých rozdílech způsobených změnou konformace deoxyribosy, jako například délka jedné otáčky (jedenáct párů bází). Tato struktura existuje při nedostatečné solvataci, což je typické pro krystaly. K jejímu objevu vedla právě snaha o krystalizaci DNA.[3] Další formou je levotočivá šroubovice, která se vyskytuje v oblastech DNA, kde se střídá jedna purinová báze s pyrimidinovou. To umožní překlopení struktury a uspořádání pentosa-fosfátové kostry do struktury připomínající písmeno Z, podle něhož se nazývá Z-DNA. DNA se však v buňkách nachází v různých formách nadšroubovicového uspořádání, které neuniklo pozornosti vědců snažících se vysvětlit stabilizaci těchto struktur. Roku 1959 Karst Hoogsteen pozoroval pomocí rentgenostrukturní analýzy netypické párování mezi adeninem a thyminem a mezi guaninem a protonovaným cytosinem. Ukázalo se, že tento typ párování, který byl nazván Hoogsteenův, by se v DNA, a hlavně v RNA, mohl vyskytovat poměrně často a rozšiřovat tak její strukturní a funkční diverzitu.[4] Později byly objeveny úplně nové strukturní motivy uspořádání do trojšroubovice, ve kterých hraje významnou roli právě vodíkový kation, který protonuje purinové/pyrimidinové zbytky nukleotidů. Podle toho byla tato forma DNA nazvána H formou.[5]

RNA

RNA má v organismu celou škálu funkcí a s nimi spojené struktury. Na rozdíl od DNA nedochází při její syntéze k formování dvoušroubovice. Je syntetizována jako samostatné vlákno, ke kterému není dosyntetizováno druhé komplementární vlákno. Může však nabývat různého prostorového uspořádání, ovlivněného jak WC (Watson-Crick), tak Hoogsteenovým párováním. Známé jsou například struktury jetelového listu u tRNA obsahující vazebné místo pro aminokyseliny a antikodonovou sekvenci, která rozpozná sekvenci na mRNA, která slouží jako templát pro proteosyntézu.

V souvislosti s proteosyntézou byl objeven další typ párování, který popsal Francis Crick při pokusu objasnit fenomén známý jako degenerace genetického kódu. Fakt, že se v DNA nachází 4 báze a kodony mRNA se skládají ze 3 bází, umožňuje existenci 64 kodonů. Standardní genetický kód používá 61 kodonů pro 20 aminokyselin a 3 stop kodony pro signalizaci terminace translace. Transferových RNA, které přenášejí tyto aminokyseliny, je však ve většině organismů méně než 45.[6] To znamená, že jedna tRNA musí být schopná se komplementárně párovat s více kodony. To Crick vysvětlil tak, že první báze na 5' konci antikodonu tRNA nemá vyhraněnou jen jednu komplementární strukturu pro tvorbu páru. Právě typ partnera na mRNA určuje geometrii báze a celkovou geometrii párování tRNA.[7] Podle konformační volnosti umožňující malé změny konformace antikodonové sekvence, byl tento typ párování nazván „wobble“. Brzy se ukázalo, že wobble párování hraje úlohu u všech typů ribonukleových kyselin a jejich biologických funkcí.[8] Alternativy párování ale nejsou jediným faktorem, který umožňuje diverzitu struktur RNA. Molekuly RNA navíc obsahují velké množství modifikovaných nukleosidů, jakými jsou například inosin a pseudouridin. Díky rozmanitosti struktur mají tyto molekuly velmi široké spektrum biologických funkcí. Mnohé z nich, jako například enzymatická aktivita nebo regulace genové exprese, budou v práci dále probírány, jelikož v jejich chemické podstatě hrají velkou roli právě tautomerní rovnováhy či protonace a deprotonace bází jednotlivých nukleotidů.

1.2 Modifikace složek nukleových kyselin

Modifikace nukleotidů a nukleosidů je jeden z nejsložitějších nástrojů evoluce. Právě toto vylepšení struktur totiž zabezpečuje různé důležité biologické vlastnosti, jako například rezistenci hub či mikrobů proti virům potřebnou k jejich přežití. Navíc tyto látky představují velký potenciál v medicíně například při léčbě virových či nádorových onemocnění.

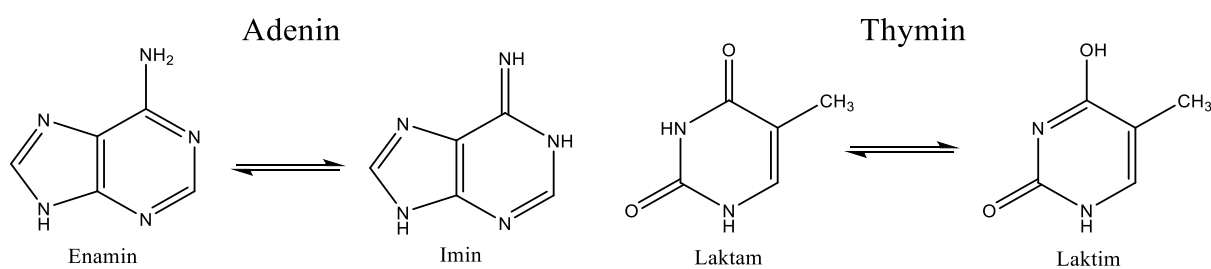
1.2.1 Modifikace chemické

K modifikacím může docházet na purinovém/pyrimidinovém skeletu, stejně tak jako na cukerné složce nebo na fosfátovém řetězci. Co se týče cukerného zbytku, dochází hlavně k methylaci ribosy na hydroxylu v pozici 2. Význam této modifikace je zatím předmětem výzkumu. Předpokládá se, že jednou z funkcí je inhibice poyladenylace.[9] Různé uměle připravené deriváty, které vznikají substitucí atomů nebo celých funkčních skupin na ribosovém kruhu, mají široké uplatnění při léčbě virových či nádorových onemocnění.[10] Mezi takové sloučeniny patří acyklické fosfonáty nukleosidů, které mají fosfátovou skupinu připojenou přes methylenový můstek zajišťující větší metabolickou stabilitu. Jedním z nich je Tenofovir, který je nejčastěji užívanou látkou na zabránění reprodukce viru HIV v těle. Za jeho syntézou stojí známý český vědec Antonín Holý.[11] Modifikace fosfodiesterové kostry se používá především ve farmaceutickém průmyslu a genetickém inženýrství na ochranu DNA a RNA před degradací nukleasami. Jednou z nejznámějších metod je tvorba fosforothioátů nahrazením kyslíku atomem síry.[12] Zajímavé je, že teprve nedávno byla stejná metoda „ochrany“ DNA před štěpením dvoušroubovice objevena u bakterií.[13] Nejpestřejší je škála modifikací poslední složky, pyrimidinových/purinových bází, která zahrnuje reakce jako je methylace, hydroxymethylace, redukce, oxidace, transglykosylace, izomerace atd. K těmto modifikacím dochází hlavně u RNA, která tak může vykonávat velké množství rozdílných biologických funkcí.[14]

1.2.2 Modifikace fyzikálně-chemické

Tautomerie nukleových bází

Díky možnosti protonace a deprotonace atomů dusíku a karbonylových skupin mohou báze nukleových kyselin existovat v různých tautomerních formách. I když volné báze nukleových kyselin mají mnohem větší škálu tautomerních forem než ty, které jsou vázané v nukleových kyselinách a mají „uzamknutý“ dusík v pozici 9 u purinů a 1 u pyrimidinů, dochází i v těchto makromolekulách k izomeraci. Adenin a cytosin přecházejí nejčastěji mezi amino a imino (enamin a imin) formou, zatímco guanin, thymin a uracil mezi keto a enol (laktam a laktim) formou (Obr. 1). Na vlastnosti nukleových kyselin mají takovéto změny struktury silný vliv. Aby vznikaly v DNA WC páry bází, musí tyto složky existovat ve svých keto a amino formách. Ionizovatelné skupiny bází mají hodnoty pK_a daleko od hodnoty fyziologického pH ~ 7 ($pK_a = 4.2$ pro N3 cytidinu, $pK_a = 3.5$ pro N1 adenosinu $pK_a = 9.2$ pro N1 guanosinu anebo $pK_a = 9.2$ pro N3 uridinu).[15] Amino a keto formy proto převažují nad minoritními formami o faktor 10^4 až 10^5 . Rozdíl v termodynamické stabilitě převažujících a minoritních forem je však přibližně 23-29 kJ/mol. Tato hodnota naznačuje, že na tautomerní rovnováhy můžou mít velký vliv mnohé faktory, jako například pH, teplota nebo přítomnost iontů kovů a enzymů.[16]



Obr. 1: Příklady tautomerních forem adeninu a thyminu

DNA

Je zjevné, že při migraci vodíkového protonu po atomech dusíku a karbonylových skupinách purinu nebo pyrimidinu se výrazně mění jejich donor-akceptorové vlastnosti, a to vede ke změnám v párování vodíkovými můstky. Tento fakt naznačuje, že by právě přítomnost různých tautomerů mohla mít za následek mutace při replikaci DNA.[17, 18] Dlouho byly mutace v DNA způsobené přítomností různých tautomerů bází jen předmětem

teoretických studií nebo výpočtů. Nedávno však byly tyto hypotézy potvrzeny, například u mechanismu mutagenních vlastností 5-aza-5,6-dihydro-2'-deoxycytidinu, což je syntetický nukleosid, který právě pomocí stabilizace svých tautomerních forem při fyziologickém pH způsobuje mutace při replikaci viru HIV, a tím inhibuje jeho replikaci,[19] nebo při objasnění mechanismu chybovosti DNA polymerasy I u organismu *Bacillus Stearothermophyllus*. U tohoto organismu se ukázalo, že jeho polymerasa někdy sama stabilizuje nesprávnou tautomerní formu cytosinu.[20]

RNA

Na rozdíl od DNA, hrají u RNA právě minoritní tautomerní formy bází významnou roli v rozšiřování strukturní a funkční diverzity těchto molekul. Jak bylo zmíněno výše, tautomerní rovnováhy jsou v blízkosti fyziologického pH výrazně posunuty ke keto a amino formě, což je způsobené hodnotami pK_a jejich ionizovatelných skupin. Pomocí fluorescenčních substituentů byly objeveny posuny hodnot pK_a až o 4 jednotky k neutrálnímu pH způsobené stabilizací sekundárními a terciárními strukturami nukleových kyselin.[21] Takovéto posuny mohou výrazně ovlivnit rovnovážné poměry jednotlivých forem. Tento fakt výrazně podporuje předpoklad účasti tautomerie a ionizace v mechanismech více biologických funkcí RNA. Mezi ně patří například enzymatická aktivita ribonukleových enzymů tzv. „ribozymů“. Jedná se o třídu malých samoštěpících enzymů, jejichž minoritní tautomerní formy bází se účastní acidobazické katalýzy autolytického štěpení fosfodiesterové kostry RNA.[22] Další funkcí je regulace genové exprese tzv. RNA přepínači (z angl. „riboswitch“), což jsou netranslatované součásti mRNA. Tyto přepínače obsahují malou doménu nazývanou aptamer, která specificky váže určité molekuly jako signály pro změnu konformace, a může tak ovlivňovat úroveň translace proteinu. V neposlední řadě bylo dokázáno, že modifikace tRNA na uridinu, které stabilizují minoritní tautomerní formy bází, hrají roli v degeneraci genetického kódu.[23]

1.2.3 Protonace a deprotonace

Tautomerizace je rychlý dynamický proces, při kterém se protonuje jedna ionizovatelná skupina molekuly a deprotonuje se jiná část. Protonace, deprotonace a tautomerní rovnováhy spolu tedy úzce souvisí. Jak již bylo řečeno, pK_a ionizovatelných skupin může být strukturně výrazně posunuté k fyziologickým hodnotám pH. Účast

ionizovaných forem byla prokázána ve více biologických systémech. Studium acidobazických rovnováh a struktur ionizovaných forem je důležité pro lepší pochopení jejich biologické aktivity. Znalost mechanismů, jakými tyto systémy fungují, může být přínosná pro oblast jejich potenciálního využití v medicíně nebo v biotechnologickém průmyslu.

1.3 Vlivy ionizace bází na strukturu nukleových kyselin

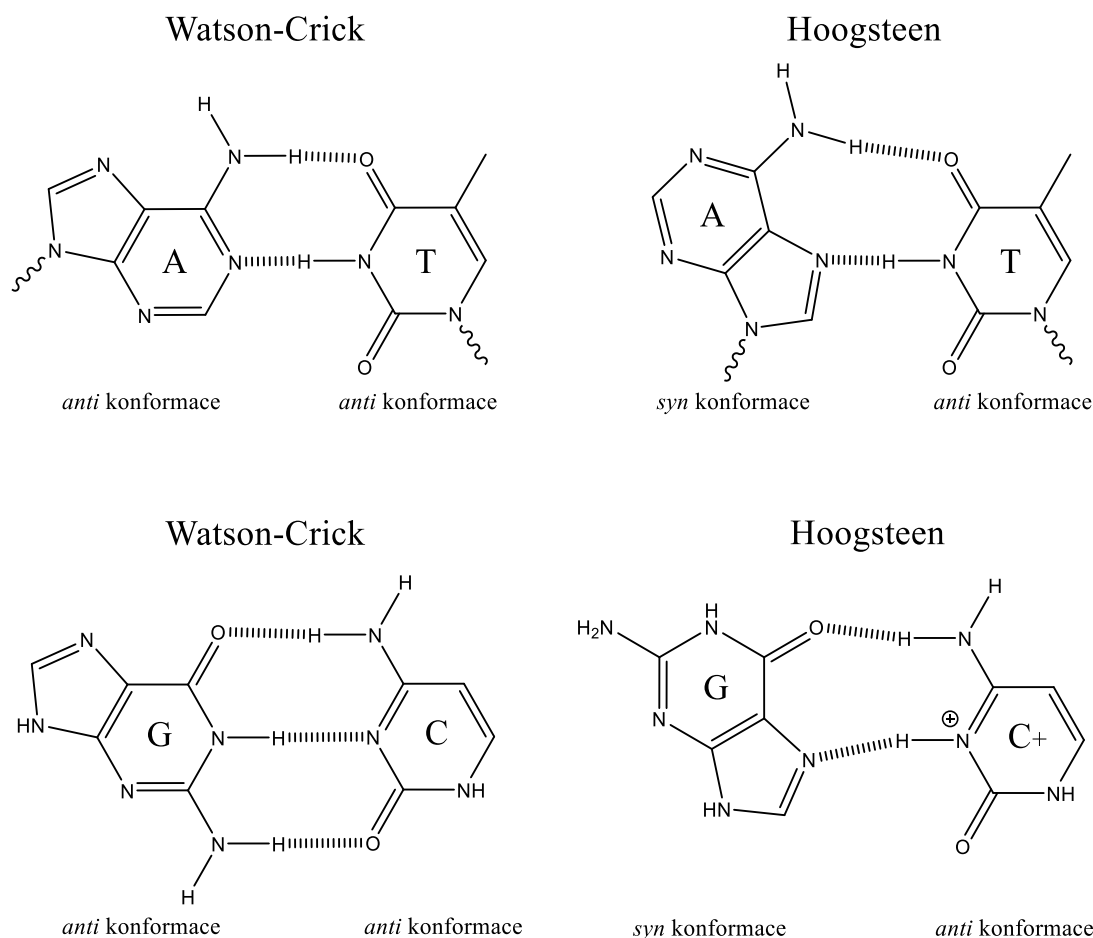
1.3.1 pH prostředí jako evoluční faktor

Zajímavou otázkou je, proč právě adenin, thymín, cytosin, guanin a uracil plní funkci uchovávání a replikace genetické informace a proč nebyl tento základní systém evolucí skoro vůbec pozměněný. Albert Eschenmoser a jeho spolupracovníci se pokusili tento fakt vysvětlit a provedli sérii experimentů, ve kterých postupně nahrazovali fosfodiesterovou kostru a báze jinými potenciálně přírodními složkami, které by mohly vést ke vzniku organizovaných struktur.[24] Pozorovali neočekávané trendy v párování a stabilitě duplexů, které se pokusili vysvětlit silou elektrostatických interakcí, keto-enol tautomerií a strukturou skeletu. Po neúspěšných pokusech dokázat vliv těchto faktorů na stabilitu duplexu, objevili korelaci jiné fyzikálně-chemické vlastnosti se schopností párování. Touto vlastností bylo pK_a jednotlivých bází. Po hlubší analýze tak vyvodili závěr, že pokud je pK_a heterocyklu blízké pH ($pK_a - pH < 2$) média, ve kterém se nachází, dochází k jeho ionizaci. Ionizovaná forma se stává hydrofilnější, zvýší se její afinita k okolní vodě a drasticky se sníží její schopnost interakce s komplementární bází. Naopak, pokud je rozdíl pK_a a pH média dostatečně velký ($pK_a - pH > 2$), báze, jakožto hydrofobní molekuly, mají tendenci agregovat, a tím je stabilita jednotlivých komplementárních párů posilněna.[25] Právě pK_a jednotlivých kanonických nukleových bází tedy může být zásadní vlastností, která je činí vhodnými pro uchovávání genetické informace.[26]

Někdy však může být ionizace požadovaným jevem. Existují případy, kdy dochází k posunu pK_a báze blíže k neutrálním hodnotám různými chemickými modifikacemi, změnou sekundární anebo terciární struktury (hlavně u RNA) nebo konjugací s kationty kovů.[27, 28] Výsledkem toho je, že báze může existovat v ionizovaném stavu či v různých tautomerních formách při fyziologickém pH. Dnes známe velké množství případů, kde ionizované formy a minoritní tautomery nukleových bází hrají roli v katalytické aktivitě ribozymů nebo ve strukturní diverzitě pramenících z alternativních modelů párování.

1.3.2 Hoogsteenův model

Brzy poté, co Watson a Crick publikovali jejich strukturní model DNA, se po světě rozběhla vlna pokusů o bližší a přesnější popis dimerů purinových a pyrimidinových bází pomocí rentgenostrukturní analýzy. Jedním z těchto pokusů je i experiment Karsta Hoogsteena, který se snažil popsat strukturu kokrystalů obsahujících 9-methyladenin a 1-methylthymín. Methylové skupiny zde byly použity na blokování vzniku vodíkových můstků v místech, kterými se báze v DNA vážou na cukerný zbytek. Pozorovaný však byl alternativní způsob tvorby vodíkových můstků. Místo typického WC modelu s můstky mezi N1 adeninu - N3 thyminu a NH₂ adeninu - O4 thyminu byla pozorována struktura, kde byl jeden adenin otočený opačně a s N3 thyminu se pároval N7 adeninu (Obr. 2, str. 16).[29] V DNA je tato geometrie umožněna otočením adeninu o 180° okolo glykosidické vazby (N9 – C1') a změnou jeho konformace z *anti* na *syn*. Hoogsteen předpokládal, že ke stejné konformační změně může dojít i u cytosinu. V tomto případě by však N3 cytosinu musel být protonovaný (C⁺), aby mohlo dojít k tvorbě vodíkového můstku (Obr. 2, str. 16). Tento předpoklad se potvrdil roku 1968 pomocí studie optické rotační disperze DNA při různých hodnotách pH.[30] Potvrzení možnosti existence těchto struktur vysvětlilo, jak dochází k formování několika již objevených struktur, mezi které patří hlavně trojšroubovicové formy nukleových kyselin.[31] Přestože existence C⁺G párů by vysvětlovala mnohé struktury, jejich existence *in vivo*, tedy při fyziologickém pH, se nezdala pravděpodobná z důvodu nízké hodnoty pK_a N3 cytosinu ve volných nukleotidech (pK_a = 4,2).[32] Podrobné studie trojšroubovicových struktur však dokázaly, že pK_a N3 cytosinu může být výrazně posunutá a protonovaná forma stabilizována lokálním okolím stejně tak jako vodíkovým můstkem s N7 guaninu. Byly pozorovány posuny pK_a až o 5 jednotek oproti volnému nukleotidu.[33] Navíc další studie naznačují, že tyto struktury jsou velmi senzitivní na změny pH, z čehož vyplývá, že právě protonace cytosinu hraje klíčovou roli ve stabilitě, a tím i v jejich biologické funkci.[34]



Obr. 2: Watsonův-Crickův a Hoogsteenův model párování DNA bází

1.3.3 Triplexy stabilizované Hoogsteenovým párováním

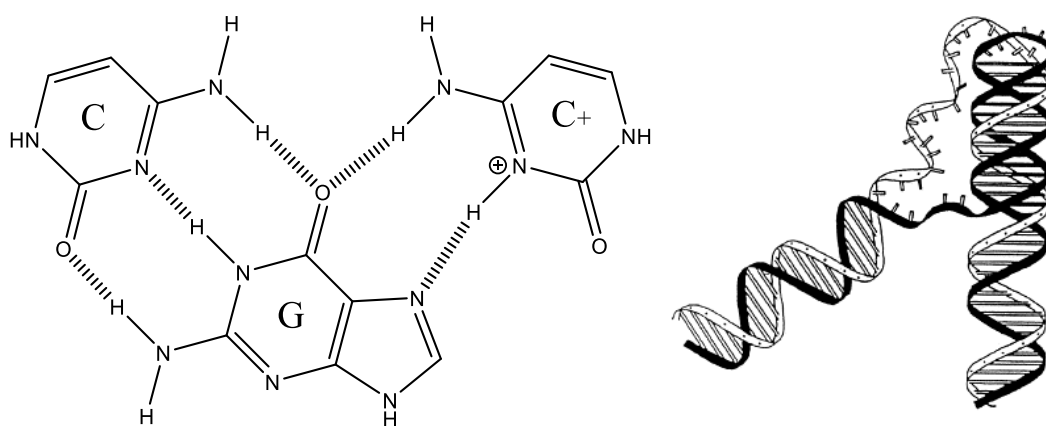
Jak již bylo výše zmíněno, Hoogsteenův model párování bází je spjatý hlavně s existencí trojšroubovicových struktur. Vznik trojšroubovicové struktury je podmíněný sekvencí nukleotidů na jednotlivých vláknech nukleových kyselin. Pokud je jedno vlákno složeno převážně z nukleotidů obsahující pyrimidinové báze (homopyrimidinové vlákno) spojené WC páry v dvoušroubovici s vláknem obsahujícím velké množství purinových bází (homopurinové vlákno), vzniká ve velkém žlábkou dvoušroubovice potenciální místo pro spárování s třetím vláknem pomocí Hoogsteenova párování, ve kterém hraje důležitou roli právě protonace cytosinu.[35]

Pokud se takto spáruje dvoušroubovice s dalším oligonukleotidovým vláknem, dochází k tvorbě trojšroubovice. Podle původu třetího vlákna je dělíme na dvě skupiny: intramolekulární a intermolekulární. Intermolekulárně vznikají tak, že se na dvoušroubovicovou DNA připojí třetí vlákno pocházející z jiné molekuly oligonukleotidu.[35] Intermolekulární formy jsou zajímavé hlavně proto, že přípravou

správného komplementárního řetězce tzv. TFO (z angl. „triplex forming oligonucleotide“) je možné cíleně ovlivňovat expresi určitých genů.[36] Tato tzv. „antisense“ technologie je dnes velmi populární genovou terapií.

Intramolekulární triplexy, nazývané H-DNA právě podle protonovaného cytosinu,[5] vznikají specifickým prostorovým uspořádáním dvoušroubovice tvořené homopurinovým a homopyrimidinovým vláknem. K jejich tvorbě dochází hlavně v oblastech zrcadlově se opakujících sekvencí zvaných H-palindromy, které byly takto nazvané právě podle H-formy DNA.[37] Ohybem dvoušroubovicové struktury ve středu symetrie a vytvořením druhé série vodíkových můstků homopyrimidinového vlákna s částí homopurinového vlákna pomocí Hoogsteenova párování se utvoří triplex (Obr. 3). Část jednoho z vláken zůstává nespárována. Existence těchto triplexů potvrzuje hypersenzitivita sekvencí na štěpení S1 nukleasami, které štěpí právě jednovláknové zbytky nukleových kyselin.[38]

Jelikož bylo prostorové uspořádání homopurinových/homopyrimidinových sekvencí do trojšroubovice dokázáno při fyziologických podmínkách *in vitro*, [5, 39] předpokládá se, že tyto sekvence tvoří triplexy také *in vivo*. Vysoký a nenáhodný výskyt těchto sekvencí v mnohých eukaryotních genomech naznačuje, že můžou hrát úlohu v regulaci biologických procesů, jakými jsou například replikace DNA.[40]



Obr. 3: Vznik intramolekulárního triplexu DNA

Vlevo: Tvorba druhé série vodíkových můstků guaninu s protonovaným cytosinem třetího vlákna

Vpravo: Formování intramolekulárního triplexu. Převzato z [41].

1.3.4 Další struktury ionizovaných bází

I když triplexy stabilizované C^+G páry pravděpodobně hrají více regulačních rolí v organismu,[5] zdaleka nejsou jedinými strukturami, jejichž vznik umožňuje existenci ionizovaných bází při fyziologickém pH. Mezi další pozorované případy patří protonované cytosiny a adeniny samoštěpících RNA enzymů – ribozymů, jejichž katalytická aktivita hraje významnou roli v procesech jako je sestřih RNA nebo replikace virů.[42, 43] V těchto případech může dokonce docházet k posunu pK_a více způsoby než jen vytvořením páru pomocí vodíkového můstku s jinou bází.[15] A^+C páry patří mezi stabilizační prvky více RNA struktur, jako například některých antikodonových smyček důležitých pro funkci těchto molekul.[44, 45] Existence ionizovaných forem a s nimi spojených struktur zároveň poskytuje vysvětlení pro různé poruchy. Například C^+C homodimery umožňují vznik tetraplexových struktur nazývaných DNA i-motivy, které jsou spojovány s genetickými poruchami jako syndrom fragilního X chromozomu nebo kosterní dysplazie.[46, 47]

Krom toho, že tyto motivy byly pozorovány v souvislosti s biologickou aktivitou nukleových kyselin, je předmětem výzkumu i potenciální využití v medicíně či biotechnologii. Například mechanismus účinku některých antibiotik je založen na jejich navázání na sekvenci DNA prostřednictvím C^+G párů.[48] Protonace guaninů a cytosinů, která způsobuje v závislosti na malých změnách pH změny sekundárních struktur biopolymerů, má naproti tomu velký potenciál pro nanotechnologie, jakými jsou nanosenzory nebo cílená distribuce léčiv v organismu.[49, 50]

1.4 Ionizované a minoritní tautomerní formy v biologických procesech

1.4.1 H-DNA a regulace genové exprese

Jak už bylo řečeno, homopurinové/homopyrimidinové sekvence DNA, které formují intramolekulární triplexy využívající protonovanou formu cytosinů, mají vysoké zastoupení v mnohých eukaryotních genomech.[40] Byly provedeny bioinformatické analýzy prokaryotických i eukaryotických genomů za účelem přesnějšího určení počtu a hlavně lokace dlouhých homopurinových sekvencí (15 – 30 purinů) a velmi dlouhých homopurinových sekvencí (>30 purinů). Statistická analýza genomů eukaryotních

organismů (kvasinky, háďátko, octomilka, myš a člověk) poukázala na velmi vysoké zastoupení 15 – 30 purinových sekvencí ve všech případech. Velmi dlouhé homopurinové sekvence byly pozorovány hlavně u myši a člověka. Naopak u prokaryot podobné homopurínové sekvence nebyly pozorovány.[51] Hlubší analýza celého genomu *S. cerevisiae* s úmyslem najít H-palindromy s délkou větší než 12 párů bází, které mají potenciál vytvářet stabilní triplexy za fyziologických podmínek, přinesla důležité výsledky. Ze 148 potenciálních sekvencí, které byly nalezeny, se nacházelo 71 z nich v oblasti genových promotorů. Spolu s faktem, že některé oblasti eukaryotních genových promotorů vykazují hypersenzitivitu na S1 nukleasy, tyto výsledky poukazují na regulační funkci triplexových struktur při genové expresi.[52]

Proto bylo navrženo velké množství experimentů založených na vkládání podobných sekvencí do oblastí genových promotorů a na zkoumání zvýšení/snížení genové exprese.[52] Mezi prvními byla například exprese genu pro β -laktamasu po vložení triplex-vytvářející sekvence do různých oblastí v okolí promotoru u *E. Coli*. Vložení sekvence před promotor vedlo k výraznému zvýšení exprese,[53] zatímco vložení mezi promotor a kódující sekvenci vedlo k silné inhibici exprese.[54] Podobné experimenty vedoucí k podobným výsledkům byly brzy provedeny i na myších a kvasinkových buněčných liniích.[55, 56]

Mezi zkoumané eukaryotní geny, jejichž promotory obsahují homopurinové/homopyrimidinové sekvence, patří například geny kódující slepičí malát dehydrogenasu (dekarboxylující),[57] lidský gama-globin,[58] lidský c-ets-2,[59] myší alfa-aktin vaskulárního hladkého svalstva,[60] myší c-ki-ras,[61] lidský decorin[62] a další. Funkční analýzy těchto sekvencí *in vitro* odhalily, že všechny mají tendenci tvořit intramolekulární triplexy a hrají významnou roli v regulaci transkripční aktivity. Mechanismy, kterými tyto struktury ovlivňují transkripci, jsou zatím předmětem hypotéz. Vznik H-DNA by například mohl vést k rozvinutí promotorových sekvencí umožňujících tak začátek transkripce. Specifické prostorové uspořádání oblasti promotoru nebo vznik jednovláknového úseku, podmíněný vznikem intramolekulárního triplexu, může zas ovlivňovat vazbu různých regulačních proteinů, jelikož je známo několik H-formujících sekvencí v oblastech promotorů, na které se regulační proteiny vážou.[63] Je však nutné připomenout, že existují i studie, které zpochybňují vliv intramolekulárních triplexů na regulaci transkripce. Například některé mutační analýzy genu C1 inhibitoru naznačují, že i přes existenci homopurinové/homopyrimidinové sekvence v jeho promotoru nemá tvorba triplexu skoro žádný vliv na transkripci.[64] Tento fakt potvrzuje skutečnost, že regulace

transkripce intramolekulárními triplexy DNA a mechanismy, jakými tyto procesy probíhají, jsou stále neprobádanou oblastí, která vyžaduje další studium a vývoj nových metod, jakou je například spolehlivá identifikace místa protonace bází nukleových kyselin.

1.4.2 H-DNA a regulace replikace DNA

H-DNA struktury byly studovány i v souvislosti s replikací DNA. Mechanismus, jakým intramolekulární triplexy pravděpodobně ovlivňují replikace, je takzvané „uvěznění“ DNA polymerasy a zastavení pohybu replikační vidlice. Ideálním příkladem jsou DNA polymerasy, které jsou schopné rozplést DNA bez přítomnosti helikasy. Netemplátové vlákno tak zůstává volné, a když polymerasa narazí na H-palindrom, volné vlákno se spáruje s dvoušroubovicí a vytvoří triplex, který zastaví replikaci. Navíc tak zabrání možnosti replikace na daném volném vlákně. Aby mohl být tento mechanismus dokázán, byly uměle pozměněny sekvence v oblasti H-palindromů polymerasy T7 tak, aby snížily potenciál těchto sekvencí tvořit triplexy, a zabránily tak terminaci replikace. U mutací obnovujících tento potenciál byla terminace znovu pozorována.[65] Stejně výsledky byly pozorovány pro Vent DNA polymerasu a DNA polymerasu I, ale ne pro Taq polymerasu, která je známá tím, že endonukleolyticky štěpí netemplátové zbytkové vlákno, takže nedochází k vytváření triplexu a k terminaci replikace.[66]

Vznik intramolekulárních triplexů ale není podmíněný prvním rozpletením DNA dvoušroubovice. Tyto struktury mohou v oblastech homopurinových/homopyrimidinových sekvencí vznikat spontánně a ukázalo se, že mají potenciál inhibovat také postup některých DNA helikas, jakou je například velký T-antigen viru SV40 (z angl. „Simian vacuolating virus 40“).[67] Inhibice replikace H-DNA byla pozorována *in vivo*, právě při studii vlivu $(dG-dA)_n \cdot (dT-dC)_n$ sekvence na terminaci replikace DNA viru SV40.[68]

Uspořádání DNA do H-formy může tedy zabránit její replikaci jak inhibicí syntézy nascentního vlákna DNA polymerasami, tak inhibicí rozplétání dvoušroubovice helikasami nebo polymerasami. Existuje velké množství studií zkoumajících regulaci replikace homopurinovými/homopyrimidinovými úseky DNA, protože tato metoda regulace představuje velký potenciál například pro boj proti replikaci virů v organismu.[40]

1.4.3 RNA přepínače

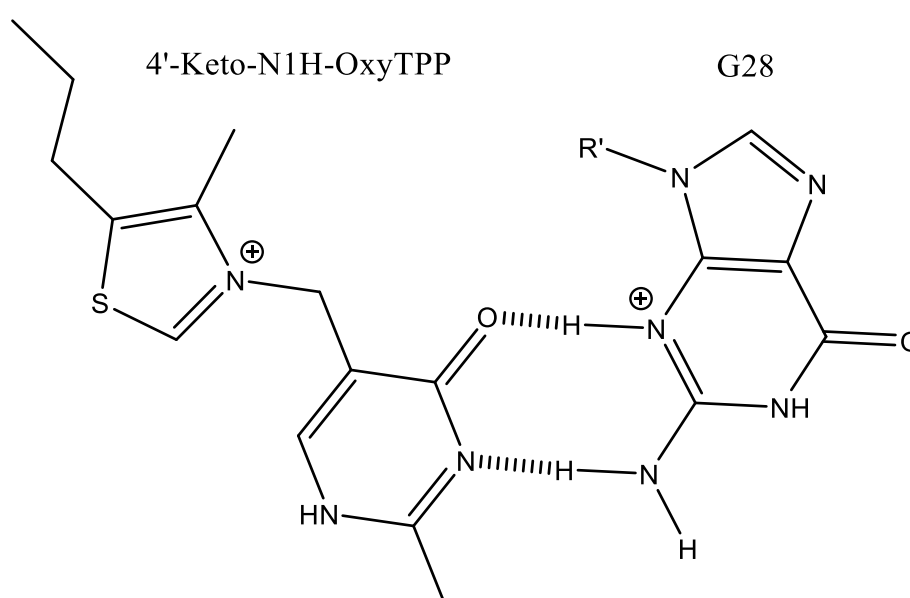
Mezi další z pozorovaných systémů regulace genové exprese, ve kterém hraje roli protonace a existence minoritních tautomerních forem nukleových bází, patří tzv. RNA přepínače. Jsou to molekuly mRNA, které jsou schopné regulovat expresi sekvence, kterou samotné nesou. Obsahují totiž 2 strukturní domény: aptamer[69] a expresní platformu.[70] Aptamer je schopný vázat různé ligandy fungující jako regulační faktory, protože po jejich navázání změní mRNA konformaci i v oblasti expresní platformy, čímž ovlivní expresi určitého proteinu.[71] Vazba ligandu je však přísně podmíněná strukturou jak aptameru, tak ligandu samotného. Jelikož se jednotlivé tautomerní nebo ionizované formy bází od sebe liší polohou donorů i akceptorů vodíkových vazeb, může být pro vazbu ligand – aptamer nevyhnutnou podmínkou právě existence určité specifické formy.

V této souvislosti byly pozorovány dva RNA přepínače: purinový přepínač, který se účastní regulace metabolismu adeninu a guaninu[72] a thiaminpyrofosfátový (TPP) přepínač hrající roli v negativní regulaci expresi genů potřebných pro biosyntézu a transport tohoto koenzymu.[70] Jelikož ve většině případů funguje regulace exprese na principu inhibice produktem,[71] byly v obou případech zkoumány právě interakce s molekulami podobnými produktům metabolických drah, které tyto RNA přepínače regulují.

V případě purinového přepínače se ukázalo, že interaguje taktéž s purinovými metabolity, a to konkrétně s hypoxantinem, xantinem a syntetickými analogy, i přestože na základě strukturních analýz tyto interakce nebyly předpokládány. Za specifickou vazbu ligandu jsou totiž zodpovědné zbytky C74, U51, U47 a U11, které vytvářejí vazebné místo pro puriny s donorem vodíkové vazby v pozici 2. Xantin má ve struktuře své primární tautomerní formy v této pozici kyslík. Afinita přepínače ke xantinu by mohla být vysvětlená tím, že se xantin váže ve formě 2-enol tautomeru. Tento fakt potvrzují další výsledky, které ukazují vyšší stabilitu vazby ligandu při nižších hodnotách pH, protože enol tautomery xantinu jsou při nižších hodnotách pH stabilnější.[73]

U TPP přepínače byla vyřešena krystalická struktura v komplexech s různými ligandy, která odhalila G28 jako kritickou bázi pro specifickou vazbu s ligandy. Rentgenstrukturní analýza také ukázala, že amino skupina v pozici 4' TPP sehraje úlohu donoru vodíkového můstku na pozici N3 G28.[74] Při interakci G28 s různými ligandy byla pozorována úplně stejná interakce s OxyTPP, potenciálním metabolitem TPP, který však má v této pozici keto skupinu.[75] Ta ale není schopná fungovat jako donor. Proto byly provedeny další

experimenty s deriváty OxyTPP, aby se zjistilo, zda se tento ligand neváže ve formě enol tautomeru. Výsledky však tuto teorii vyvrátily. Přítomnost velmi podobných interakcí, v případě obou ligandů TPP a OxyTPP s RNA přepínačem, tedy naznačuje, že N3 G28 je při interakci s OxyTPP protonovaný a funguje jako H-donor (Obr. 4).[76] I když výsledky jak purinového, tak TPP přepínače ukazují na existenci uvedených tautomerů a protonovaných forem, nebyly tyto formy přímo pozorované, a neexistuje tak experimentální důkaz jejich existence, a to z důvodu nedostatku metod schopných pozorování tak rychlých procesů, jako je například výměna protonů ve vodném médiu. To opět poukazuje na důležitost metod schopných těchto pozorování, a tedy na význam této práce.



Obr. 4: Vazba OxyTPP na G28 TPP RNA přepínač

4'-Keto-N1H-OxyTPP, potenciální metabolit TPP, interaguje s aptamerem TPP RNA přepínače pravděpodobně prostřednictvím 2 vodíkových můstků s protonovaným G28.

1.4.4 RNA katalýza

Navzdory omezené komplexitě stavebních jednotek RNA (4 kanonické báze) jsou tyto molekuly schopné efektivně vykonávat celou škálu biologických funkcí. Jednou z nich je intenzivně zkoumaná schopnost katalyzovat některé biochemické procesy. Ukazuje se, že ionizace nukleovýchází hraje důležitou roli u RNA katalýzy, umožňující RNA enzymům (ribozymům) katalyzovat reakce stejně efektivně jako proteinovým enzymům.[42]

Předpokládá se, že tyto katalytické procesy by mohly využívat 2 třídy ionizovaných struktur: třída I zahrnuje ionizované báze, které jsou spárované s jinými bázemi, zatímco

třída II se těchto interakcí neúčastní.[15, 77] U spárovaných ionizovaných nukleobází I. třídy se předpokládá účast na katalýze vytvořením tzv. oxyaniontové díry, struktury, která stabilizuje negativní náboj deprotonovaného kyslíku nebo alkoxidu v tranzitním stavu.[78] U peptidových enzymů hraje tuto úlohu obvykle lysin nebo arginin. V případě, že se ionizovaná forma párování neúčastní (třída II), může sehrát roli přímo v acidobazické katalýze, podobně jako histidin u proteinových enzymů.[15]

V souvislosti s ionizovanými formami bází je intenzivně studováno 5 přirozeně se vyskytujících ribozymů: vlásenkový (z angl. „hairpin“) ribozym, ribozym typu hammerhead, ribozym netranslatované oblasti glukosamin-6-fosfát synthetasy (glmS), ribozym viru Hepatitis Delta (HDV) a Varkud satellite ribozym (VS). Navzdory tomu, že mechanismus jejich katalýzy je téměř shodný, strukturně i svým původem se výrazně liší. Hammerhead a vlásenku můžeme najít ve virech napadajících rostliny,[79] VS v mitochondrii různých hub [80] a HDV zase ve viru lidského patogenu.[81] GlmS funguje zároveň jako RNA přepínač v regulaci produkce glukosaminu-6-fosfátu u Gram pozitivních bakterií.[82]

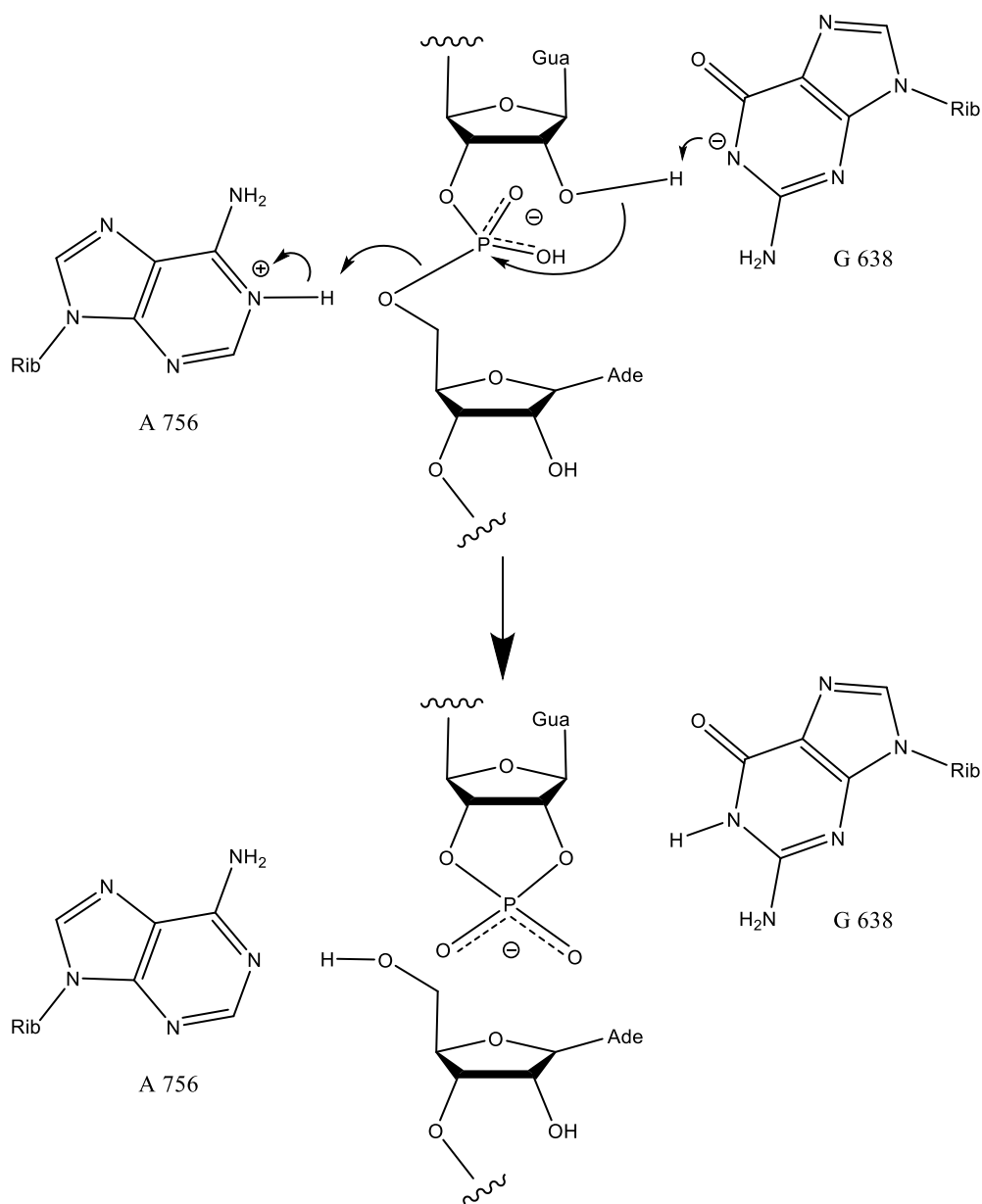
U všech těchto enzymů se předpokládá podobný mechanismus katalýzy jako u RNAsy A.[22] Katalytická schopnost štěpu fosfodiesterové vazby enzymu spočívá ve 4 krocích: správném prostorovém uspořádání nukleofilu, štěpitelného fosfátu a odstupující skupiny pomocí uspořádání nukleotidů, aktivaci 2'-OH nukleofilu bází, zvýšením afinity k tranzitnímu stavu a aktivací 5'-O odstupující skupiny kyselinou.[83] U ribozymů je studována role katalytických guanosinů, u kterých se předpokládá, že právě ionizace a tautomerie hrají roli v aktivaci 2'-OH nukleofilu. Guanosiny s deprotonovaným dusíkem N1 mají potenciál fungovat jako báze, které jsou schopné aktivovat 2'-hydroxyl a zapříčinit tak jeho nukleofilní atak na vedlejší 3'-fosfát. Tato vnitřní transesterifikační reakce vede k vytvoření 2 separovaných molekul RNA.[22] Strukturní a biochemické studie ribozymu Hammerhead, vlásenkového ribozymu a glmS odhalily, že N1 katalytických guanosinů se ve všech případech nachází v blízké vzdálenosti (do 3.5 Å) 2'-hydroxylu sousedícího se štěpeným 3'-fosfátem.[84, 85] U VS ribozymu byla krystalická struktura vyřešena poměrně nedávno (2013) a výsledky rovněž ukazují stejný strukturní prvek v aktivním místě tohoto ribozymu (G638) (Obr. 5, str. 25).[86] V případě HDV tuto roli sehrává Mg^{2+} v hydratované formě $[(MgOH)]^+$, který je nezbytný pro funkci tohto ribozymu.[42]

I přesto, že důležitost guanosinů při aktivaci 2'-hydroxylu byla potvrzena, hodnoty pK_a těchto skupin nejsou v souladu s navrhovaným mechanismem. pK_a N1 guanosinu se

pohybuje kolem hodnoty 10 a při fyziologickém pH by měl být plně protonován, a být tedy chemicky nevhodný pro deprotonaci 2'-hydroxyly, jehož pK_a je přibližně 13.[87] Deprotonovaná forma N1 je ale nukleofilní a je schopna aktivace hydroxylové funkční skupiny. Jak již bylo řečeno, hodnoty pK_a nukleobází mohou být výrazně posunuté v závislosti na blízkém okolí, sekundární struktuře, přítomnosti iontů atd.

Celý autolytický proces štěpení RNA však potřebuje také protonaci 5'-O odstupující skupiny kyselinou. Z pěti zmíněných ribozymů je HDV ideálním příkladem katalýzy kyselou skupinou, kde tuto roli sehrává protonovaný N3 a C75.[88] U vlásenkového ribozymu tuto roli zastává protonovaný N1 a A38.[43] Poslední z ribozymů o kterém se předpokládá, že má funkci protonované báze jako donoru protonu 5'-O skupině, je VS. Jedná se o protonovanou formu jeho A756 (Obr. 5, str. 25).[80] Hammerhead k této funkci využívá pravděpodobně 2'-OH skupinu G8[89] a glmS využívá dokonce C2-aminoskupinu cukerného zbytku GlcN6P (glukosamin-6-fosfát).[90]

Všechna uvedená pozorování předpokládají účast ionizovaných forem bází v acidobazické katalýze, tedy existenci II. třídy ionizovaných struktur. Objevují se však studie, které naznačují účast ionizovaných forem I. třídy na katalýze elektrostatickou stabilizací elektronegativního tranzitního stavu. V souvislosti s existencí tohoto alternativního mechanismu je studován vlásenkový ribozym,[91] ale například i 23S podjednotka ribozymu u *E. Coli*, konkrétně A2452 jeho rRNA.[92]



Obr. 5: Mechanismus acidobazické katalýzy VS ribozymu

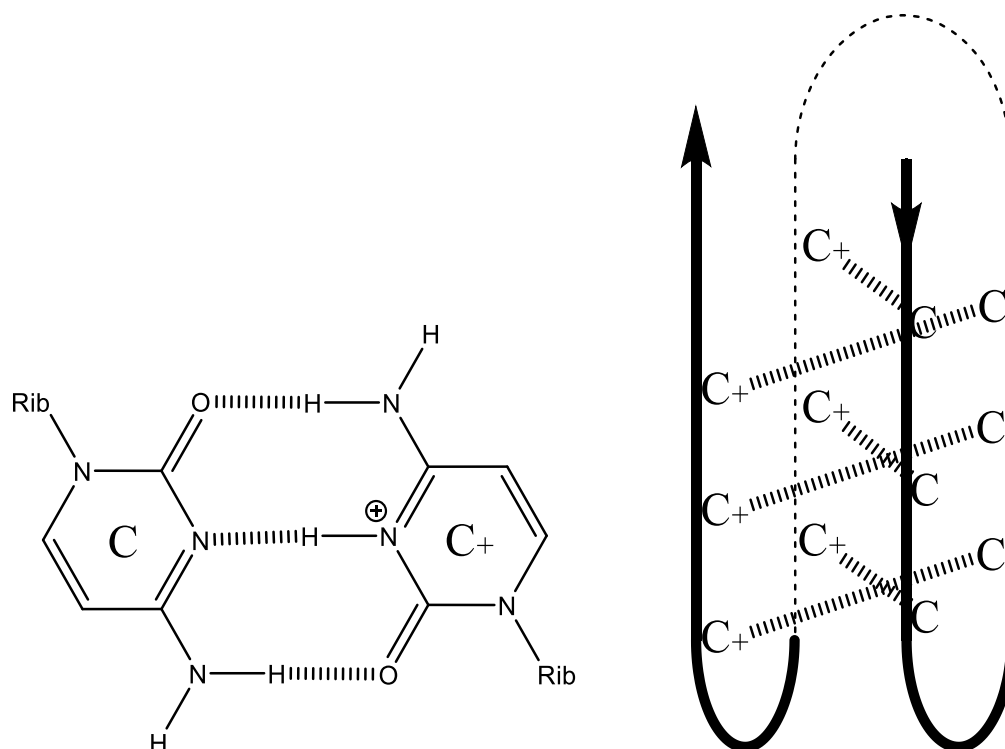
Deprotonovaný G638 fungující jako báze poskytuje volný elektronový pár vodíku 2' hydroxylu. Kyslík se stává nukleofilním, dochází k S_N2 nukleofilní substituci na 3'-P. 5'-O, odstupující skupina je následně protonována na hydroxyl protonovanou formou A756. Výsledkem je rozštěpená fosfodiesterová vazba na cyklický 2'3'-fosfát a 5'-hydroxyl.

1.4.5 Poruchy spojené s ionizovanými formami

Pozměněné rozložení elektrostatického potenciálu u ionizovaných forem bází jim umožňuje tvořit páry, které nejsou v souladu s běžným WC modelem. Důsledkem existence těchto párů je vznik několika sekundárních struktur DNA a RNA, jakými jsou již zmíněné triplexy stabilizované C^+G Hoogsteenovy páry nebo tetraplexy. Ty můžou vznikat

vytvořením vodíkových můstků mezi cytosiny dvou vlásenek (z angl. „hairpin“) (Obr. 6, str. 27). Podmínkou však je, aby byla polovina cytosinů protonována, a mohla tak fungovat jako donory ve vodíkové vazbě.[46] Přítomnost těchto sekundárních struktur může způsobovat různé problémy při replikaci, mezi které patří expanze repetitivních trinukleotidových úseků.[93] K expanzi dochází při replikaci, pokud DNA polymerasa narazí na úsek templátového vlákna, který obsahuje tyto opakované sekvence, a po syntéze určité části komplementárního vlákna dojde k „vyklouznutí“ DNA polymerasy. Nově nasyntetizované vlákno je uvolněno a dochází k vytvoření strukturního motivu vlásenky, která vzniká spárováním bází stejného vlákna mezi sebou, přičemž dojde k odhalení jedenkrát zreplikovaných bází templátového vlákna. DNA polymerasa zůstává přítomna v místě, kde došlo k pozastavení replikace, a začíná znovu syntetizovat komplementární vlákno, a to na konci vytvořené vlásenky. V důsledku toho danou sekvenci nasyntetizuje vícekrát, čímž dojde k její multiplikaci.[94] DNA polymerasy jsou však vysoce specifické a efektivní enzymy a k uvolnění syntetizovaného vlákna nedochází bezdůvodně. K takovým chybám dochází v oblastech trinukleotidových opakujících se sekvencí kvůli existenci triplexů a tetraplexů, které mohou bránit polymerase v pohybu a ovlivňovat tak správnost její funkce.[95] Výsledkem může být expanze repetitivních úseků, která vede k dalšímu posilnění tvorby těchto struktur, a tím k inhibici transkripce specifického proteinu. Tímto případem je Friedreichova ataxie, která je zapříčiněna sníženou expresí proteinu frataxinu. Snížená exprese tohoto proteinu ve spojitosti s neobvyklými DNA strukturami byla pozorována v buňkách ledvin afrických zelených opic.[96]

Buňky ovšem mají mnohé mechanismy na opravu chyb vzniklých při replikaci a takto vzniklé vlásenky jsou často vystříhnuty kontrolními enzymy, protože jsou nestabilní a obsahují nesprávný počet spárovaných bází. Tady ale přichází do hry tetraplexové struktury, které dokáží vlásenky mezi sebou vytvářet a tím stabilizovat svoji strukturu a zabezpečit si existenci do dalšího cyklu replikace (Obr. 6, str. 27). Při analýzách těchto útvarů, spojených například se syndromem fragilního chromozomu X nebo kosterní dysplazií, byly několikrát pozorovány posuny pK_a cytosinů až k hodnotám vyšším než 8. Tyto hodnoty dokazují jejich protonaci při fyziologickém pH a tedy existenci C^+C párů.[46, 47] Mimo zmíněný syndrom křehkého chromozomu X (expanze CGG tripletů) a kosterní dysplazie (expanze GAC tripletů), patří mezi choroby zapříčiněné expanzí repetitivních trinukleotidových úseků například Friedreichova ataxie (expanze GAA/TTC tripletů) nebo Huntingtonova choroba (expanze CAG tripletů).[97]



Obr. 6: model C⁺C páru a příklad struktury DNA tetraplexu (i-motiv)

Protonace N3 jednoho z cytosinů umožňuje vznik stabilního páru se 3 vodíkovými můstky. Tento stabilizační prvek umožňuje vznik tetraplexu, v tomto případě i-motivu, který může vytvářet například 2 vlásenky vznikající blízko sebe při chybné replikaci.

1.4.6 Potenciální terapeutické a biotechnologické využití

Účast ionizovaných či tautomerních forem bází nukleových kyselin v biologických procesech a existence s nimi spojených struktur přináší velký potenciál jak pro léčbu různých onemocnění, tak pro biotechnologie nebo genové inženýrství. V současnosti jsou intenzivně zkoumané oligonukleotidy (TFO), které jsou schopné vytvářet triplexy se sekvenčně specifickými úseky DNA. Tyto intermolekulární triplexy vznikají vytvořením Hoogsteenových párů mezi oligonukleotidy a bázemi DNA v oblasti velkého žlábků dvoušroubovice.[98] Vznik triplexu v oblasti specifické sekvence s sebou přináší nové možnosti inhibice replikace nebo exprese různých genů, kontroly interakce DNA s proteiny, detekce specifických úseků DNA pomocí různých sond či indukce mutagenese. Metody spojené s TFO mají několik výhod oproti starším metodám, a to hlavně díky vysoké specifitě, která zabráňuje nechtěným interakcím, a malé velikosti, která umožňuje přístup k jaderné DNA přímo v buňkách.[99]

Interakce molekul s DNA využívající ionizované formy bází nemusí vést jen k tvorbě triplexu. Antibiotikum Triostin A, které má zároveň protinádorové účinky, funguje jako

interkalát DNA, který stabilizuje dvoušroubovicovou strukturu a zabraňuje přístupu polymeras. Tato stabilizace zahrnuje C^+G Hoogsteenovy páry mezi jednotlivými vlákny.[100]

Potenciál využití vlastností protonovaných cytosinů v nukleových kyselinách je předmětem zájmu i v oblastech nanotechnologie. Inspirace pro využití pochází pravděpodobně z objevů spojených s již zmíněnými genetickými poruchami, kde C^+C páry umožňují tvorbu tetraplexových struktur DNA. Jedná se hlavně o polymery schopné měnit strukturu v závislosti na pH, a tím i jejich makroskopické vlastnosti.

Mezi konkrétní testované aplikace patří například molekulové motory, tedy polynukleotidy, které v závislosti na pH okolí existují buď v kompaktní sbalené formě stabilizované C^+C páry, nebo v rozbalené formě, kdy cytosiny nejsou protonované. Tyto motory mají velký potenciál v oblasti cíleného transportu léčiv, kde je potřeba, aby byla aktivní látka vypuštěna v přesně definovaném místě. To může být zajištěno právě díky mírnému rozdílu pH.[101] Dalšími oblastmi využití těchto vlastností polynukleotidů jsou hydrogely, pH indikátory nebo inteligentní povrchy.[102]

2 Cíle práce

Acidobazické a tautomerní rovnováhy nukleových kyselin ovlivňují v mnoha směrech biologickou aktivitu nukleových kyselin a jejich složek. Tato bakalářská práce má dva cíle. Prvním cílem je shromáždění informací týkajících se vlivu protonace/deprotonace nukleobází na biologickou aktivitu nukleových kyselin a studium případného využití těchto vlivů. Druhým cílem je vypracování metodiky vhodné pro spolehlivé určení místa protonace/deprotonace pomocí NMR (z angl. „nuclear magnetic resonance“) spektroskopie. Pro dosažení tohoto cíle byla nejprve vybrána série látek, u které je místo protonace/deprotonace dané její strukturou. U těchto látek byly při různých hodnotách pH experimentálně určeny NMR parametry. Následně byly pro tyto látky vypočteny NMR parametry pro protonovanou/deprotonovanou formu a vypočtené hodnoty byly porovnány s experimentem. Na základě výsledků byla vybrána optimální výpočetní metoda pro další experimenty. V další fázi byly naměřeny a vypočteny parametry pro složitější molekuly, kde není místo protonace jasně dáno strukturou.

3 Materiál a metody

3.1 Materiál

Sigma Aldrich, Inc. (USA)

1-methyluracil (čistota 98%); 2-amino-4-methylpyrimidin (čistota 98%); 2-amino-4,6-dimethoxypyrimidin (čistota 98%); 6-chloro-2,4-diaminomethylpyrimidin (čistota 98%); deuterovaná voda (D_2O , 99,90 atom. % D); deuterovaná kyselina chlorovodíková (DCl , 35% roztok v D_2O , 99 atom. % D); deuterovaný hydroxid sodný ($NaOD$, 40% roztok v D_2O , 99,5 atom. % D); 1,4-dioxan (dioxan, 10% roztok v D_2O); nitromethan

Penta s.r.o. (ČR)

PENTANAL kalibrační pufr borátový (koncentrovaný, pH 9); PENTANAL kalibrační pufr fosfátový (koncentrovaný, pH 7,00); PENTANAL kalibrační pufr ftalátový (koncentrovaný, pH 4,00)

3.2 Přístroje

NMR spektrometr

500 MHz Avance III vybavený kryosondou (Bruker)

pH metr

pH Probe Micro pro NMR kyvety nerezová ocel (HACH LANGE)

3.3 Software

NMR měření a analýza spekter

TopSpinTM (Bruker); MNova NMR (Mestrelab Research)

Kvantově chemické výpočty

Gaussian 09 revision D.01 (Gaussian, Inc.); GaussView 3.09 (Gaussian, Inc.)

3.4 Metody

3.4.1 Příprava roztoků bází

Přibližně 1 g 1-methyluracilu byl rozpuštěn v 10 ml D_2O . 0,5 ml tohoto roztoku bylo přeneseno do skleněné NMR kyvety. Do kyvety byla přikápnuta kapka 10% roztoku dioxanu. Po vytemperování roztoku v kyvetě na 20°C bylo změřeno a zaznamenáno jeho pH. Je nutné poznamenat, že při měření pH deuterovaného roztoku je naměřenou hodnotou pD. Ta se od pH mírně liší z důvodu izotopového efektu na měřící sondě.[103] V této práci však přesná hodnota pH není důležitá vzhledem k tomu, že pracujeme s relativními hodnotami. Proto v dalším popisu bude nadále používán výraz pH.

Následně bylo do původního roztoku 1-methyluracilu, také vytemperovaného na 20°C, přidáno asi 0,2 ml 10% dioxanu. Do roztoku byla ponořena pH sonda a za stálého míchání přikapáván 40% NaOD, zředěný D_2O v poměru 1:10, až dokud hodnota pH nestoupala o cca 0,5 až 1,0. Přesná hodnota pH byla zaznamenána a z roztoku bylo opět odebráno 0,5 ml do NMR kyvety. Takto bylo získáno celkem 10 roztoků v celé škále bazických hodnot pH.

V případě 2-amino-4-methylpyrimidinu byl postup stejný, přikapávána byla 35% DCl, zředěná D_2O v poměru 1:10. Hodnota pH tedy postupně klesala a do NMR kyvet byly odebrány roztoky po změně pH v rozmezí přibližně 0,5 až 1,0. Takto bylo získáno celkem 9 roztoků v celé škále kyselých hodnot pH.

U 2-amino-4,6-dimethoxypyrimidinu byl postup stejný jako v předešlém případě, získáno bylo 8 roztoků ve škále kyselých hodnot pH.

V případě 6-chloro-2,4-diaminomethylpyrimidinu bylo rozpuštěno cca 0,1 g v 0,5 ml D_2O přímo v NMR kyvetě a po vytemperování roztoku na 20°C a přidání kapky 10% dioxanu bylo změřeno a zaznamenáno jeho pH. Následně byl stejným způsobem připraven ještě jeden roztok, který byl okyselen pár kapkami 35% DCl až do silně kyselého roztoku, a jeho pH bylo opět zaznamenáno.

3.4.2 NMR spektroskopie

NMR spektroskopie je jednou z nejvyžívanějších metod na zkoumání jak struktury organických molekul, tak dynamických procesů, kterých se tyto molekuly účastní. Je zde využita specifická vlastnost jader izotopů prvků, které se běžně vyskytují v organických molekulách. Jedná se především o izotopy ^1H a ^{13}C . Touto vlastností je takzvaný spin umožňující jejich detekci v silném magnetickém poli. V experimentu na zkoumání jádra ^1H a ^{13}C byla pozorovaná změna jejich stínění elektrony v závislosti na pH. Zkoumanou veličinou byl chemický posun δ , který je na intenzitě stínění závislý.

V programu TopSpinTM byla nastavena pulzní sekvence měření jednotlivých jader. NMR sonda byla nejprve naladěna na rezonanční frekvenci měřeného jádra. Následně bylo provedeno frekvenční uzamknutí signálu rozpouštědla D_2O , tzv. „lock“ a automaticky byla naladěna homogenita magnetického pole. U všech roztoků byly měřeny experimenty ^1H ; ^{13}C APT (z angl. „attached proton test“); ^1H , ^{13}C HMBC (z angl. „heteronuclear multiple bond correlation“) a pro každou bázi byly u neutrálního a nejbazičtějšiho/nejkyselějšího roztoku změřeny ^1H , ^{15}N HMBC. Chemické posuny ve spektrech ^1H NMR a ^{13}C NMR byly vztaženy na standard 1,4-dioxan a posuny v ^1H , ^{15}N HMBC NMR byly vztaženy na standard nitromethan. Chemické posuny jednotlivých standardů:

1,4-dioxan: ^{13}C NMR (126 MHz, D_2O): $\delta = 67,19 \text{ ppm}$. ^1H NMR (500 MHz, D_2O): $\delta = 3,75 \text{ ppm}$.

nitromethan: ^{15}N NMR (51 MHz, D_2O): $\delta = 381,7 \text{ ppm}$.

3.4.3 Optimalizace geometrie a výpočet NMR parametrů

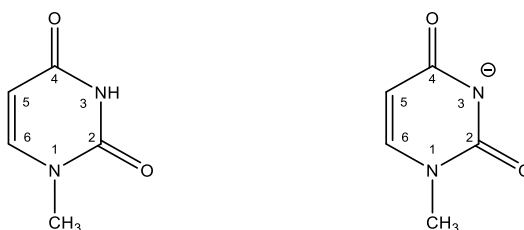
Výpočetní chemie se stává stále více používaným nástrojem pro lepší pochopení či odhad některých chemických dějů. Pomocí kvantově chemických výpočtů je možné vytvořit simulaci struktur nebo různých vlastností molekul, které jsou velmi blízké realitě. V této práci je používanou metodou DFT (z angl. „density functional theory“), která je založena na předpokladu, že energie a vlastnosti mnohočásticových systémů jsou jednoznačně určeny jejich elektronovou hustotou. Tato metoda byla použita na optimalizaci struktur derivátů pyrimidinovýchází a výpočet jejich NMR parametrů. Zkoumané byly vypočtené stínící konstanty. Jejich hodnoty jsou závislé na elektronové hustotě v okolí jednotlivých jader, které stíní vnější magnetické pole, a tím definují chemický posun jádra v NMR spektru. Všechny výpočty byly provedeny v programu Gaussian09, revize D.01. Jednotlivé molekuly

byly nejprve namodelovány v programu GaussView. Následně byly nastaveny parametry výpočtu: funkcionál, báze funkcí a model solvatace. Funkcionál je matematickým operátorem, báze funkcí reprezentuje atomové orbitaly použité k výpočtu molekulových orbitalů a model solvatace ve výpočtu simuluje chování vybraného rozpouštědla. Výpočet byl proveden ve výpočetním středisku ÚOCHB AVČR.

4 Výsledky a diskuze

4.1 Přiřazení signálů pomocí 1D a 2D NMR spekter

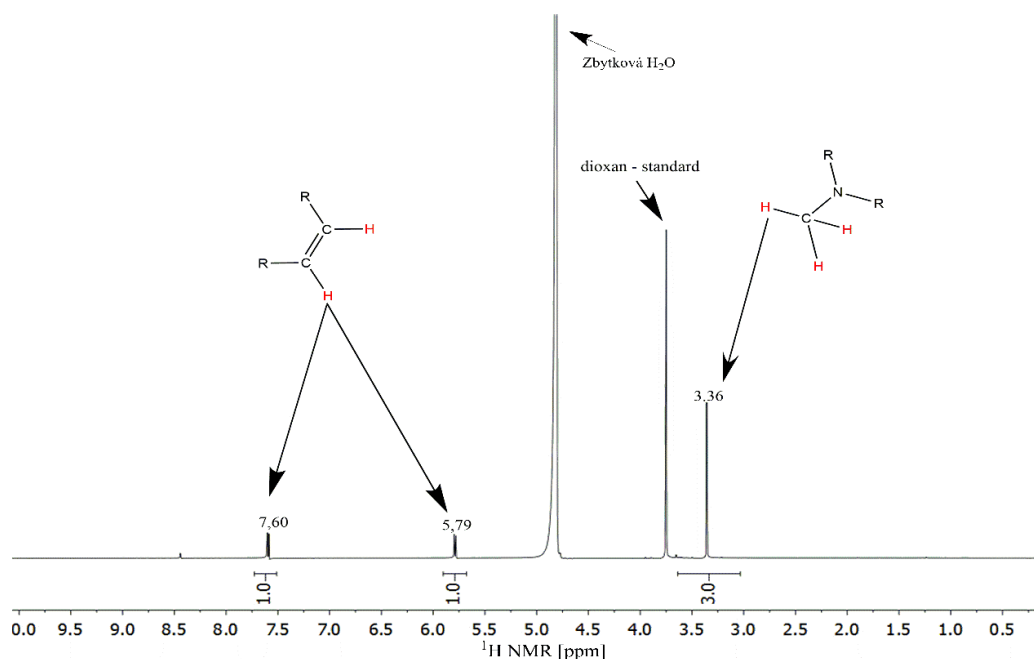
První analyzovanou látkou byl 1-methyluracil (Obr. 7). Jedná se o modelovou molekulu, protože místo deprotonace je jednoznačně definované strukturou, a to na atomu N3, čímž vzniká jeho anion (Obr. 7).



Obr. 7: 1-methyluracil (vlevo) a jeho anion (vpravo)

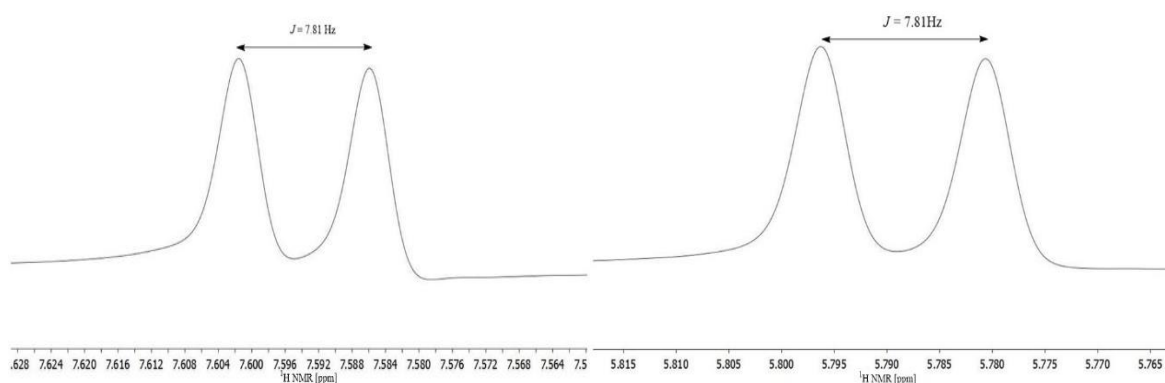
4.1.1 ^1H NMR

První krok při analýze výsledků je jednoznačné přiřazení jednotlivých atomů molekuly k signálům v NMR spektrech. Jako první je analyzované ^1H spektrum (Obr. 8).



Obr. 8: ^1H NMR spektrum 1-methyluracilu (pH 6,94)

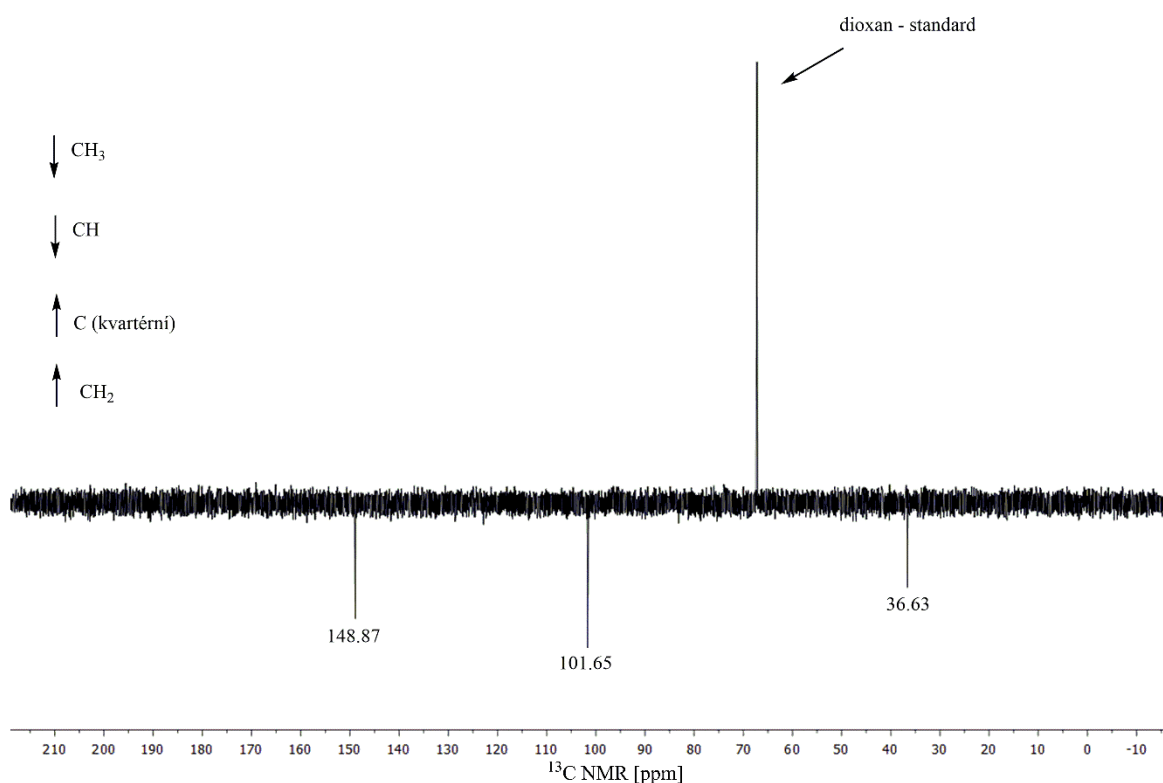
U vodíkového spektra podávají signály i kvantitativní údaje. Integrální hodnoty pod signály určují počet ekvivalentních protonů. Tři ekvivalentní protony představují jedinou methylovou skupinu, proto můžeme předpokládat, že signál 3,36 ppm reprezentuje methylovou skupinu. Druhý fakt, který je třeba vzít v potaz, je chemické složení okolí jednotlivých atomů a jeho vliv na chemický posun. Oblast od 3 do 5 ppm je typická pro vodíky vázané na uhlíku, který má vazbu na heteroatom. Oblast okolo 6 ppm je typická pro vodíky na dvojně vazbě. Tyto fakta potvrzují, že by signál 3,36 ppm mohl patřit methylové skupině. Vysoký signál u 5 ppm je způsobený zbytkovým nedeuterovaným rozpouštědlem, jedná se tedy o signál vody. Signály u 5,79 a 7,60 ppm pravděpodobně reprezentují vodíky na uhlících 5 a 6. To je možné potvrdit analýzou interakčních konstant J , které jsou také typické pro ^1H spektra. Oba signály jsou dublety a interakční konstanta je identická (7,81 Hz) (Obr. 9), což znamená, že dochází k vzájemnému štěpení a atomy se v molekule nachází blízko sebe (interakce přes 3 vazby). Signál vodíku na dusíku není ve spektru pozorovatelný, jelikož dochází k jeho záměně s deuteriem v rozpouštědle, což je velmi rychlý proces.



Obr. 9: Interakční konstanty vodíků 1-methyluracilu v ^1H NMR spektru (pH 6,94)

4.1.2 ^{13}C NMR

Druhým krokem přiřazení signálů atomům 1-methyluracilu je analýza jader ^{13}C (Obr. 10).

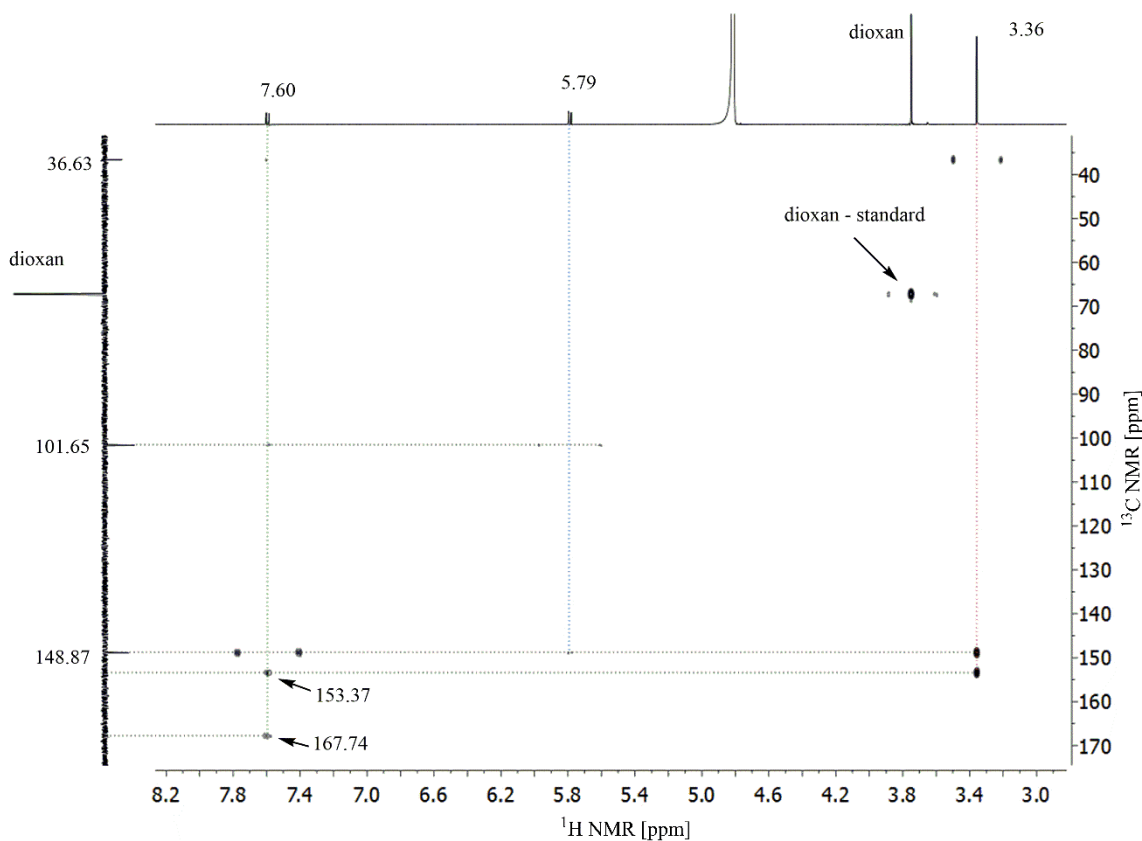


Obr. 10: ^{13}C NMR spektrum 1-methyluracilu (pH 6,94)

Jádra ^{13}C byla měřena pulzní sekvencí APT, která umožňuje okamžitě viditelné rozdělení uhlíků do 2 skupin. Signály kvartérních uhlíků a CH_2 směřují ve spektru nahoru, zatímco signály CH a CH_3 směřují dolů. Z toho jasně vyplývá, že ani jeden kvartérní uhlík (C_2 a C_4) nebyl detekovaný. Chemický posun 67,19 ppm reprezentuje standard dioxan. Signál s nejnižším posunem je pravděpodobně methylový uhlík. O tom svědčí opět charakteristická oblast posunu a fakt, že signál směřuje dolů (jedná se o CH_3). K přiřazení signálů 101,65 ppm a 148,87 ppm a detekci kvartérních uhlíků byla použita 2D NMR spektroskopie.

4.1.3 ^1H , ^{13}C HMBC NMR

^1H , ^{13}C HMBC spektrum podává informace o spin-spinových interakcích jader ^1H s jádry ^{13}C přes 2 až 3 vazby (Obr. 11).

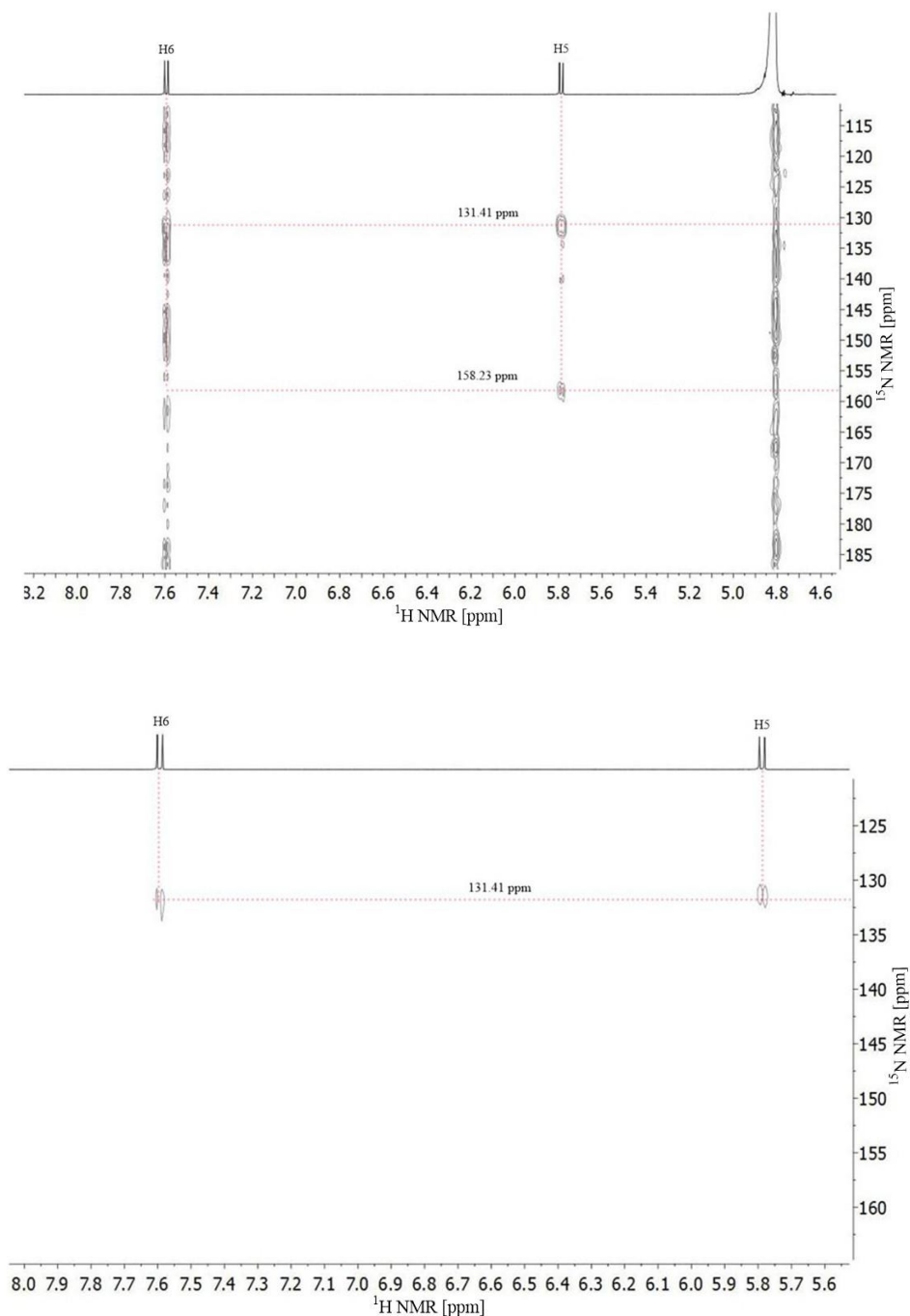


Obr. 11: ^1H , ^{13}C HMBC NMR spektrum 1-methyluracilu (pH 6,94)

Ve spektru lze pozorovat několik silných interakcí. Signál ^1H s posunem 3,36 ppm vykazuje 2 silné interakce s uhlíkovými signály, a to u 148,87 ppm a 153,53 ppm. Pokud tento signál reprezentuje methylové vodíky a signál uhlíku 148,87 ppm může být jen CH do vzdálenosti 3 vazeb, pak tento signál reprezentuje C6. Signál 153,37 ppm není v ^{13}C NMR spektru pozorovatelný, jedná se tedy o kvartérní uhlík s dosahem přes 3 vazby na methylové vodíky, což reprezentuje C2. Dalším důležitým údajem je přímá interakce přes jednu vazbu uhlíku 148,87 ppm, tedy C6 s vodíkovým signálem 7,60 ppm. Ta jasně určuje, že signál 7,60 ppm v ^1H spektru reprezentuje vodík na C6. Signál 167,74 ppm interagující s tímto vodíkem tedy reprezentuje druhý kvartérní uhlík, opět 3 vazby vzdálený C4. Signál ^1H 5,79 ppm zbývá pro vodík na C5, jehož signál můžeme pozorovat u 101,65 ppm v ^{13}C spektru (je pozorovatelná i slabá přímá interakce C5 a jeho vodíku).

4.1.4 ^1H , ^{15}N HMBC NMR

Po kompletním přiřazení uhlíkových a vodíkových signálů byly pomocí ^1H , ^{15}N HMBC experimentu získány informace o chemických posunech jednotlivých dusíků. Spektra detekují interakce jader ^{15}N s jádry ^1H přes několik vazeb (Obr. 12).



Obr. 12: ^1H , ^{15}N HMBC NMR spektrum 1-methyluracilu (pH 6,94)

Nahoře: zvýšená intenzita pro detekci signálu ^{15}N NMR 158,23 ppm. Dole: snížená intenzita pro detekci interakce signálu ^{15}N NMR 131,41 ppm se signálem H6 a potlačení šumu.

Kvůli velmi nízké citlivosti jader ^{15}N byl ve spektrech pozorovatelný velký šum. V prvním spektru je možné vidět interakce obou dusíků s vodíkem na uhlíku C5 (Obr. 12 nahoře, str. 38). Pro rozlišení jednotlivých dusíků bylo však zapotřebí snížit intenzitu signálů tak, aby byly interakce s vodíkem H6 rozlišitelné (Obr. 12 dole, str. 38). Po snížení intenzity je viditelné, že dusík s chemickým posunem 131,41 ppm interaguje s oběma vodíky, zatímco dusík s posunem 158,23 ppm interaguje jen s H5. Vzdálenost H6 od N3 je 4 vazby a od N1 2 vazby, zatímco vzdálenost H5 je stejná k oběma dusíkům, a to 3 vazby. Z toho vyplývá, že u N3 pozorujeme ve spektru 1 interakci s bližším vodíkem H5, zatímco u N1 pozorujeme interakci s oběma vodíky. Chemický posun N1 je tedy 131,41 ppm a posun N3 je 158,23 ppm.

Obdobným způsobem byly přiřazeny jednotlivé signály u všech dalších analyzovaných bází. V případě 6-chloro-2,4-diaminpyrimidinu byl však zvolen alternativní postup, který je popsán v příslušné kapitole. Kvůli velmi nízké citlivosti jader ^{15}N nebyly u žádné jiné báze než u 1-methyluracilu detekovány dusíky, a to i přes vysoký počet skenů.

Chemické posuny atomů 1-methyluracilu (pH = 6,94): ^{13}C NMR (126 MHz, D_2O): $\delta = 36,63$ (CH_3); 101,65 (C5); 148,87 (C6); 153,37 (C2); 167,74 (C4) ppm. ^1H NMR (500 MHz, D_2O): $\delta = 3,358$ (s, 3H, CH_3); 5,789 (d, 1H, H5); 7,593 (d, 1H, H6) ppm. ^{15}N NMR (51 MHz, D_2O): $\delta = 131,41$ (N1); 158,23 (N3) ppm.

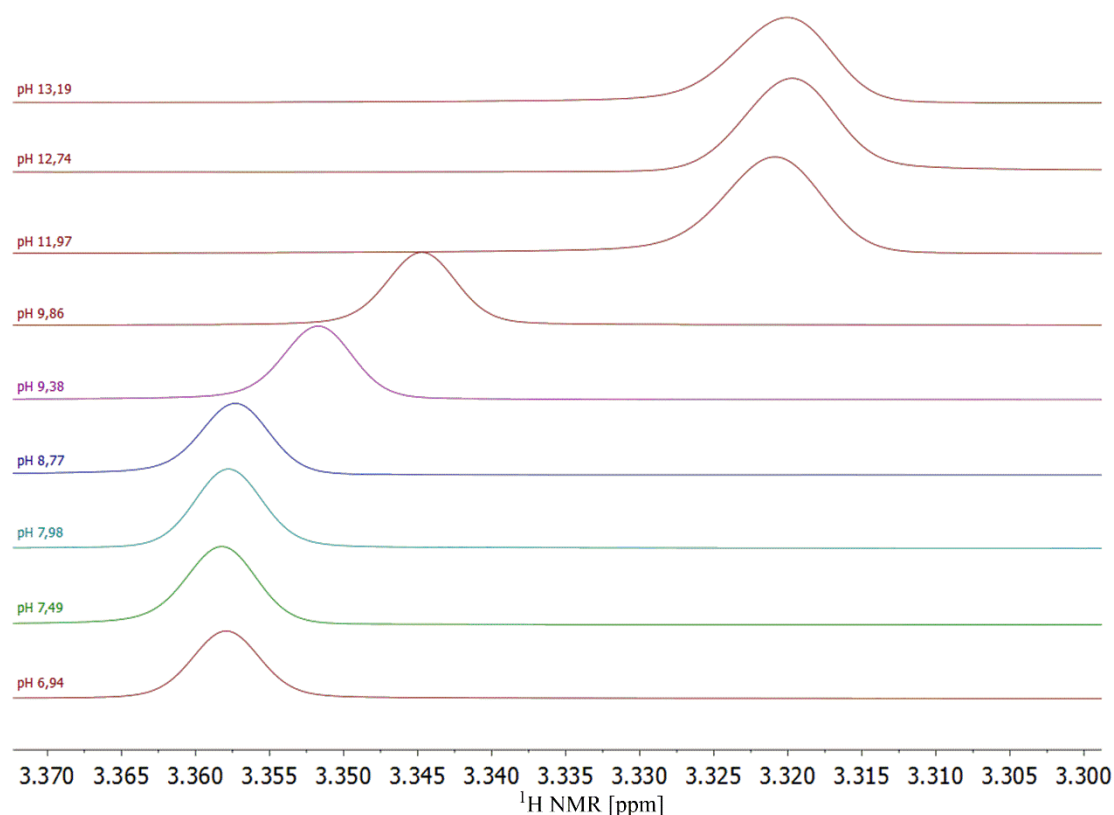
Chemické posuny atomů 2-amino-4,6-dimethoxypyrimidinu (pH = 7,00): ^{13}C NMR (126 MHz, D_2O): $\delta = 55,24$ (CH_3); 77,80 (C5); 163,03 (C2); 173,89 (C6, C4) ppm. ^1H NMR (500 MHz, D_2O): $\delta = 3,858$ (s, 6H, CH_3); 5,658 (s, 1H, H5) ppm.

Chemické posuny atomů 2-amino-4-methylpyrimidinu (pH = 6,83): ^{13}C NMR (126 MHz, D_2O): $\delta = 23,35$ (CH_3); 112,05 (C5); 158,50 (C6); 162,62 (C2); 170,28 (C4) ppm. ^1H NMR (500 MHz, D_2O): $\delta = 2,315$ (s, 3H, CH_3); 6,663 (d, 1H, H5); 8,115 (d, 1H, H6) ppm.

4.2 Titrace zásadou / kyselinou a změny chemických posunů modelových molekul

První modelová struktura, tedy 1-methyluracil, byla titrována NaOD z neutrálního do silně zásaditého pH, ve kterém by měl být N3 deprotonovaný. U každého pH byla změřena

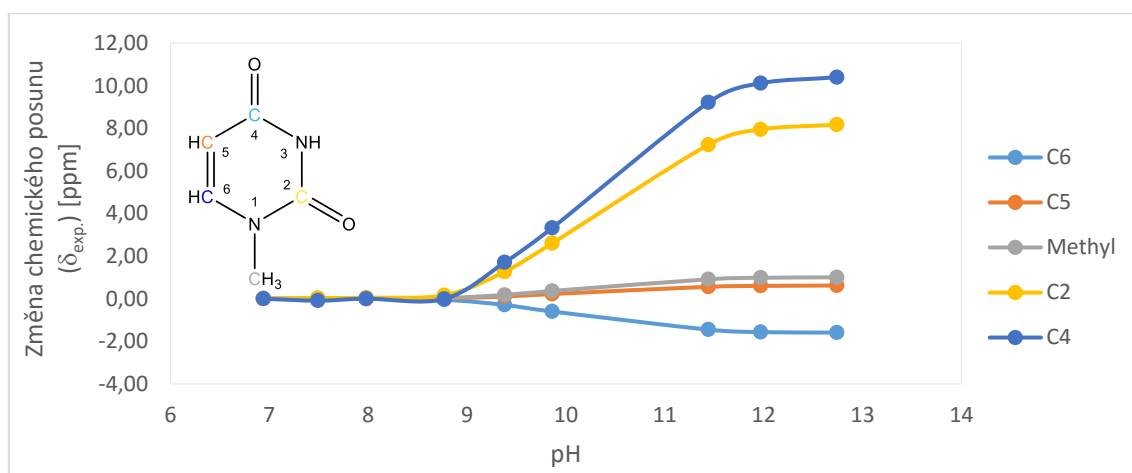
^1H a ^{13}C spektra, aby mohla být určena závislost chemického posunu jednotlivých atomů na pH. U neutrálního a nejbazičtějšího roztoku byla změřena navíc ^1H , ^{15}N HMBC spektra pro určení chemického posunu jednotlivých dusíků v protonované a deprotonované formě. Podobným způsobem byla DCl titrována druhá modelová struktura s jasně daným místem protonace, 2-amino-4,6-dimethoxypyrimidin, z neutrálního do kyselého pH. I v tomto případě byla změřena při každém pH ^1H a ^{13}C spektra a u neutrálního a nejkyselejšího i ^1H , ^{15}N -HMBC. Závislost chemického posunu na pH je možné vidět přímo ve složeném spektru 1-methyluracilu (Obr. 13).



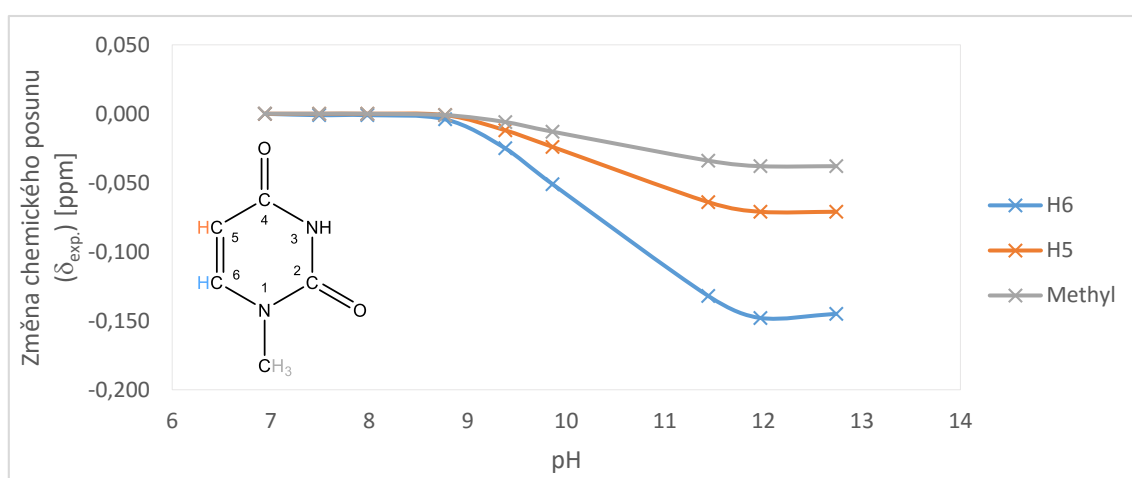
Obr. 13: Změna chemického posunu v závislosti na pH

^1H NMR signál vodíků methylové skupiny 1-methyluracilu.

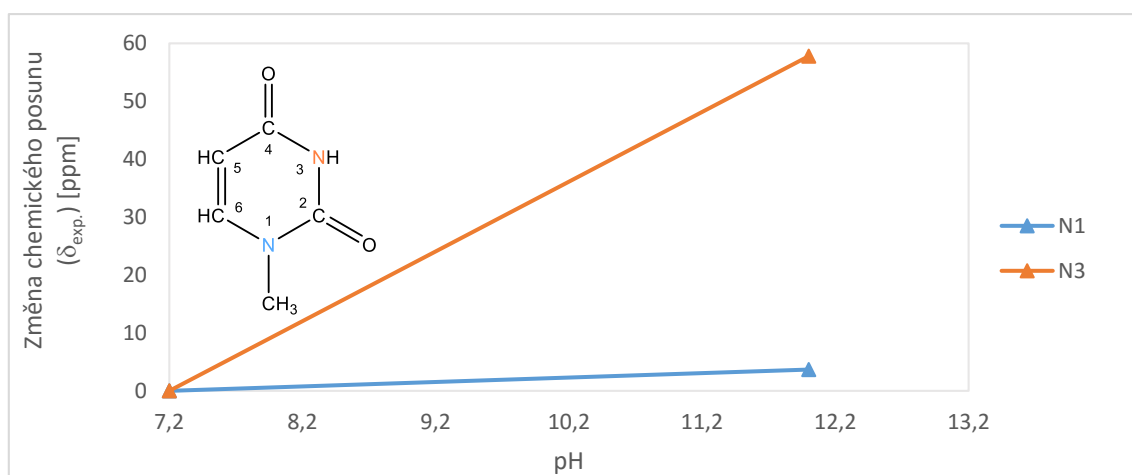
Hodnoty chemických posunů atomů při jednotlivých pH byly přepočítány na rozdíly oproti posunům v neutrálním pH ($\delta_{\text{exp.}}$). Následně byly sestrojeny titrační křivky závislosti změny chemického posunu na pH (Obr. 14, 15, 16, 17, 18 str. 41 a 42).



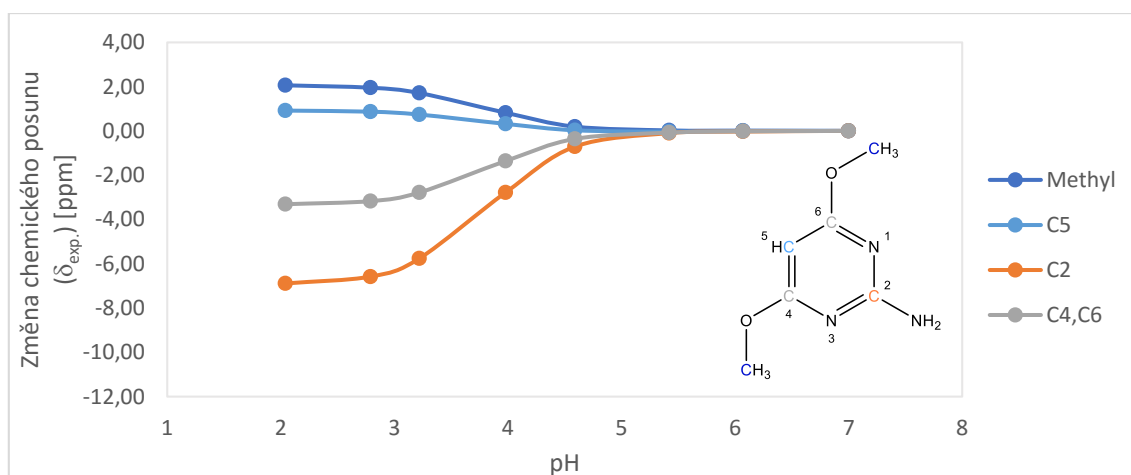
Obr. 14: Změny chemických posunů uhlíků 1-methyluracilu v závislosti na pH



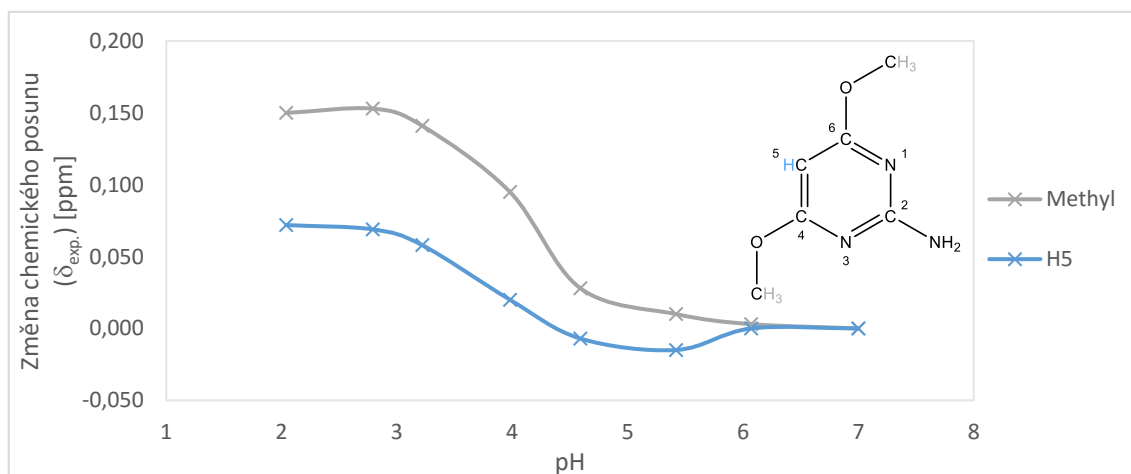
Obr. 15: Změny chemických posunů vodíků 1-methyluracilu v závislosti na pH



Obr. 16: Změny chemických posunů dusíků 1-methyluracilu v závislosti na pH



Obr. 17: Změny chemických posunů uhlíků 2-amino-4,6-dimethoxypyrimidinu v závislosti na pH



Obr. 18: Změny chemických posunů vodíků 2-amino-4,6-dimethoxypyrimidinu v závislosti na pH

Výsledné křivky potvrzují, že došlo k protonaci/deprotonaci. Navíc je zřejmé, že hodnota pH neutrálních a nejkyselejších/nejbazičtějších roztoků je dostatečně vzdálená od pK_a jednotlivých molekul. Z toho vyplývá, že v roztocích, u kterých budou experimentální NMR parametry porovnány s vypočtenými NMR parametry, se nacházejí téměř výhradně neutrální, respektive ionizované formy molekuly a ne jejich směsi. Tento fakt vylučuje, že chyby v korelaci výpočet – experiment mohou být způsobené přítomností více forem molekul v roztoku.

Na křivkách lze vidět velmi dobře srovnatelné změny chemický posunů jednotlivých atomů. V obou případech se jedná o strukturu s jasně daným místem protonace/deprotonace. U 1-methyluracilu jde o N3, zatímco u 2-amino-4,6-dimethoxypyrimidinu jde o N1 a N3, které jsou díky symetrii ekvivalentní. Proto se v kyselém roztoku nachází směs 1:1, N1

protonované a N3 protonované formy. Výrazné změny chemických posunů jsou očekávány právě v blízkém okolí ionizovatelných atomů, což titrační křivky potvrzují. Hodnoty chemických posunů uhlíků 1-methyluracilu v okolí deprotonovaného N3 se výrazně snížily, zatímco u ostatních uhlíků byla zaznamenána jen nepatrná změna. V případě 2-amino-4,6-dimethoxypyrimidinu, který se protonuje na N1 a N3, došlo naopak k výraznému kladnému posunu na uhlících C2, C4 a C6, které jsou taktéž ve vzdálenosti 1 vazby od těchto dusíků. Nižší hodnota posunu C4 a C6 je způsobena existencí dynamické rovnováhy mezi N1 a N3 protonovanými formami. Hodnota posunů se tak průměruje, zatímco na C2 má protonace na obou místech stejný vliv. Analýza vodíků poukazuje na kladné posuny při protonaci a záporné při deprotonaci, nepodává důležité informace o místě protonace, proto nebyly titrační křivky vodíků v dalších experimentech analyzované. V případě dusíků byla jádra ^{15}N detekovaná jen u 1-methyluracilu a jejich chemické posuny byly měřeny jen při neutrálním a zásaditém pH. Změna posunu N3 je výrazně větší než změna posunu N1, což potvrzuje, že došlo k deprotonaci N3. I když je zřejmé, že analýza chemických posunů dusíků má v případě zkoumání místa protonace/deprotonace největší výpovědní hodnotu, jejich problematická a časově náročná detekce tuto analýzu u dalších struktur znemožnila.

4.3 Vývoj metody detekce místa protonace korelací výpočet - experiment

Z analýzy titračních křivek je vidět, že místo protonace poměrně dobře určují změny chemických posunů na okolních uhlících. Při zkoumání struktur, u kterých však místo protonace není jednoznačně určeno strukturou, se může stát, že z experimentálních dat nebude možné vyvodit jasné závěry. Navíc již na příkladu dusíků lze vidět, že ne vždy je možné pomocí NMR spektroskopie spolehlivě detekovat všechna jádra. Také můžou struktury s několika potenciálními místy protonace existovat jako poměr jednotlivých protonovaných forem v dynamické rovnováze. V těchto případech jsou pro vyvození závěrů potřebné další informace, které můžou poskytnout právě kvantově-chemické výpočty.

4.3.1 Výběr vhodné výpočetní metody

Pro výpočet optimalizace struktury a NMR parametrů pomocí DFT výpočtů je možné volit několik typů funkcionalů aází funkcí. V této práci byl ve všech případech použit funkcional B3LYP. Pro získání co nejlepších výsledků a tedy nejmenší chyby korelace

výpočtu s naměřenými hodnotami byly parametry látky 1-methyluracilu počítané s různými bázemi funkcí a v různých simulovaných podmínkách. Použité byly báze funkcí [6-31g(p)]; [6-31g(pd)] a [6-31g(2df,2pd)]. Nejprve byla dvěma způsoby zoptimalizována struktura neutrální molekuly i aniontu. V jednom typu optimalizace byla započítaná simulace solvatace H₂O, v druhém typu proběhla optimalizace bez simulace solvatace. Pro každou optimalizovanou strukturu byly následně znovu stejnou bází funkcí vypočteny NMR parametry. V případě solvatovaných modelů byl výpočet NMR parametrů proveden se simulací solvatace i bez ní, v případě nesolvatovaných modelů byly NMR parametry spočítané bez simulace solvatace. Následně byla vypočtena hodnota rozdílu chemických posunů jednotlivých atomů mezi neutrálními a ionizovanými formami ($\Delta\delta_{\text{výpočet}}$). V dalším kroku byly hodnoty těchto rozdílů porovnány s experimentálními hodnotami rozdílů chemických posunů ionizované a neutrální formy ($\Delta\delta_{\text{exp.}}$). Výsledné hodnoty představují absolutní hodnotu odchylky naměřených a vypočtených rozdílů chemických posunů (Tab. 1). Konečným údajem je aritmetický průměr odchylek pro každý atom 1-methyluracilu. Tyto průměry odchylky experimentu od výpočtu byly stanoveny pro každou bázi funkcí a pro všechny kombinace optimalizace a výpočtu NMR (Tab. 2, str. 45)

Tab. 1: Odchylky hodnot vypočítaných a experimentálních hodnot změn chemických posunů mezi neutrální a deprotonovanou formou 1-methyluracilu

Pro výpočet použita báze funkcí [6-31g(2df,2pd)], optimalizace struktury i výpočet proběhly se simulací solvatace H₂O.

Atom C	Odchylka [ppm]	Atom H	Odchylka [ppm]	Atom N	Odchylka [ppm]
C2	0,58	H5	0,402	N1	8,16
C4	0,97	H6	0,169	N3	5,70
C5	0,23	Methyl	0,222		
C6	3,02				
Methyl	0,78				
Průměr	1,12		0,264		6,93

Tab. 2: průměrné odchylky vypočítaných a experimentálních změn chemických posunů mezi neutrální a ionizovanou formou.

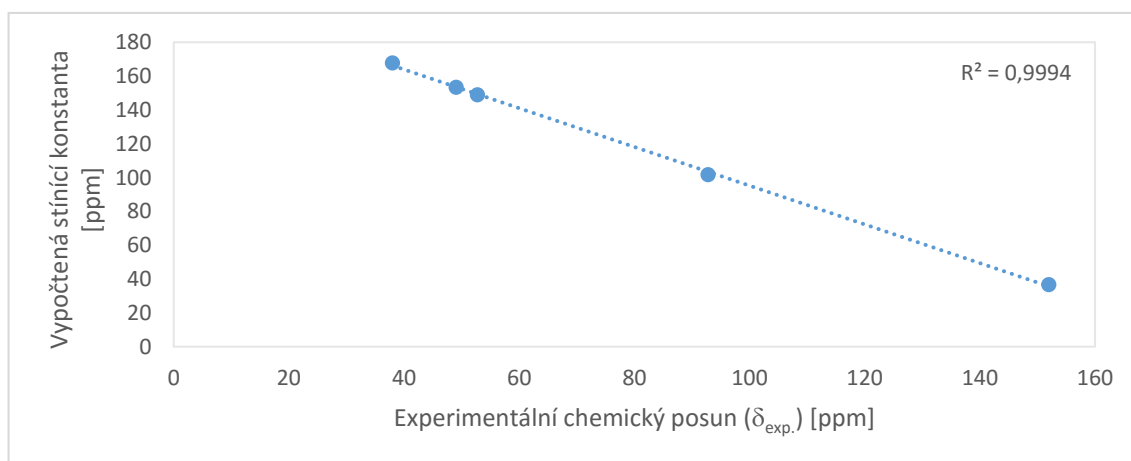
Výpočetní metoda	Opt. – vakuum NMR – vakuum [ppm]			Opt. – solvent NMR – vakuum [ppm]			Opt. – solvent NMR – solvent [ppm]		
	¹³ C	¹ H	¹⁵ N	¹³ C	¹ H	¹⁵ N	¹³ C	¹ H	¹⁵ N
B3LYP/[6-31g(p)]	2,04	0,532	5,34	1,83	0,582	9,69	1,36	0,338	7,39
B3LYP/[6-31g(pd)]	3,68	0,294	6,03	1,84	0,570	10,71	1,32	0,327	6,43
B3LYP/[6-31g(2df,2pd)]	1,96	0,424	6,28	1,82	0,475	10,62	1,12	0,264	6,93

Opt – vakuum NMR – vakuum: optimalizace struktury i výpočet NMR parametrů bez simulace solvatace; **Opt – solvent NMR – vakuum:** optimalizace struktury se simulací solvatace H₂O a výpočet NMR parametrů bez simulace solvatace; **Opt – solvent NMR – solvent:** optimalizace i výpočet NMR parametrů se simulací solvatace H₂O. **Vyznačené hodnoty:** výsledky metody s nejmenší odchylkou.

Pro výběr neoptimálnější výpočetní metody je potřebné zvážit několik faktorů a požadavky na přesnost. V případě této práce byla prioritou minimální odchylka korelace výpočtu s experimentem, a to hlavně v případě uhlíků. Vzhledem k tomu byla z jednotlivýchází funkcí vybrána pro další výpočty právě [6-31g(2df,2pd)] a optimalizace i výpočet NMR parametrů se simulací solvatace H₂O. Průměr absolutní hodnoty odchylky experimentu od výpočtu je při použití této metody ze všech možností nejnižší, a to v případě uhlíků i vodíků. Odchylka v případě dusíků není pro tuto práci až tak důležitá, jelikož v některých experimentech nebyla detekována všechna jádra ¹⁵N. Výpočet chemického posunu ¹⁵N je navíc zatížen velkou chybou z důvodů dynamických interakcí s rozpouštědlem, které je velmi složité spolehlivě nasimulovat. [6-31g(2df,2pd)] patří mezi velmi velké báze funkcí a vyžaduje spolu se simulací solvatace vysoký výpočetní výkon. Odchylky výsledků některých jednoduššíchází nejsou výrazně větší, takže pokud by byl výpočet omezený časem nebo dostupnou výpočetní technikou, což se při zkoumání složitějších biomolekul může stát, bylo by v jistých případech možné jednodušší báze funkcí použít.

4.3.2 Korelace výpočet – experiment pro ^{13}C NMR parametry neutrální molekuly

Při měření složitých molekul se může často stát, že ne všechny NMR signály bude možné jednoznačně přiřadit jednotlivým atomům. To by mohlo přinést značné komplikace při detekci místa protonace/deprotonace. Pro přiřazení jednotlivých signálů by proto mohly dobře posloužit vypočtené NMR parametry. Regresní analýzou byla analyzována korelace experimentálních chemických posunů a vypočtených stínících konstant. Protože u jader ^1H nebyla detekce či přiřazení skoro nikdy problémová, byla analyzována jen jádra ^{13}C (Obr. 19)

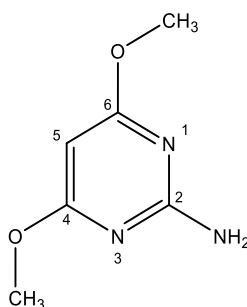


Obr. 19: Korelace výpočet – experiment pro NMR parametry ^{13}C 1-methyluracilu

Čtverec koeficientu mnohonásobné korelace (R^2), velmi blízký hodnotě 1, potvrdil přesnost výpočetní metody. Tento výsledek potvrzuje, že pomocí vypočítaných chemických posunů je možné poměrně spolehlivě přiřadit jednotlivé NMR signály.

4.3.3 Kontrola odchylky vybrané metody na látce se strukturně definovaným místem protonace

Další vybranou molekulou, jejíž struktura jasně definuje místo protonace, je 2-amino-4,6-dimethoxypyrimidin (Obr. 20, str. 47).



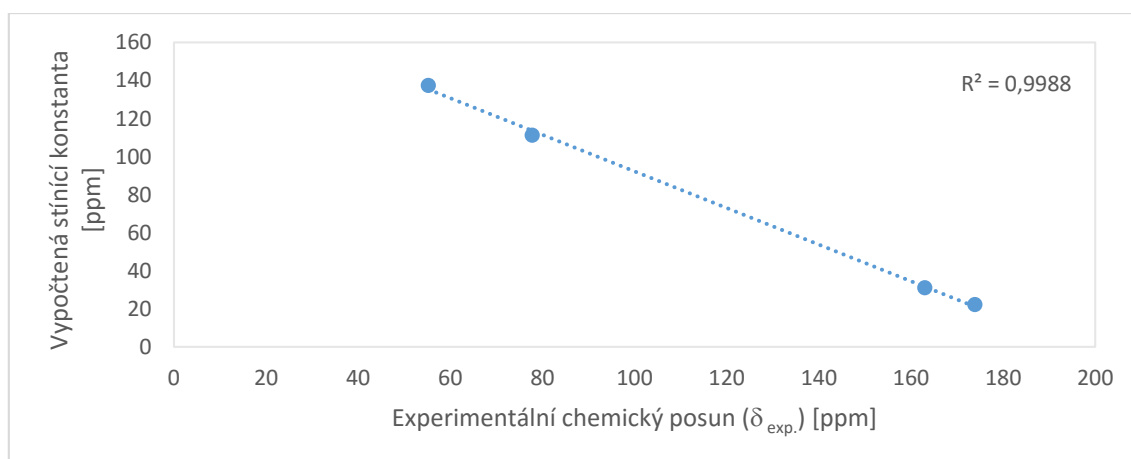
Obr. 20: 2-amino-4,6-dimethoxypyrimidin

Tato molekula je symetrická, takže dusíky N1 a N3 jsou ekvivalentní. Struktura byla zoptimalizovaná a NMR parametry vypočteny pomocí báze funkcí [6-31g(2df,2pd)] se simulací solvatace H₂O při optimalizaci struktury i při výpočtu NMR parametrů. Stejným způsobem byly vypočteny parametry pro N1 a N3 protonovanou formu a následně byly tyto hodnoty zprůměrovány 1:1, jelikož v roztoku existují tyto formy v dynamické rovnováze. Byly opět vypočteny rozdíly experimentálních hodnot posunů mezi protonovanou a neutrální molekulou. Tyto rozdíly byly porovnány s vypočtenými hodnotami (Tab. 3). Také byla analyzována korelace experimentálních chemických posunů ¹³C a vypočtených stínících konstant neutrální molekuly (Obr. 21, str. 48).

Tab. 3: Porovnání experimentálních a vypočítaných hodnot změn chemických posunů mezi neutrální a ionizovanou formou 2-amino-4,6-dimethoxypyrimidinu.

Atom	$\delta_{\text{exp.}}$ pH 7,00	$\delta_{\text{exp.}}$ pH 1,05	$\Delta\delta_{\text{exp.}}$	$\Delta\delta_{\text{výpočet}}$	Odchylka
¹³ C [ppm]					
C2	163,03	156,14	-6,89	-9,67	2,78
C4,6	173,89	170,50	-3,39	-2,59	0,80
C5	77,80	78,81	1,01	-1,04	2,05
Methyl	55,24	57,30	2,06	3,80	1,74
Průměr	-	-	-	-	1,84
¹ H [ppm]					
Methyl	3,858	4,007	0,149	0,278	0,129
H5	5,658	5,883	0,225	0,364	0,139
Průměr	-	-	-	-	0,134

$\delta_{\text{exp.}}$: experimentálně získaný chemický posun; $\Delta\delta_{\text{exp.}}$: experimentálně získaný rozdíl chemických posunů mezi neutrální a protonovanou formou molekuly; $\Delta\delta_{\text{výpočet}}$: rozdíl chemických posunů mezi neutrální a protonovanou formou molekuly získaný pomocí kvantově chemických výpočtů; odchylka: absolutní hodnota rozdílu hodnot $\Delta\delta_{\text{výpočet}}$ a $\Delta\delta_{\text{exp}}$



Obr. 21: Korelace výpočet – experiment pro NMR parametry ^{13}C 2-amino-4,6-dimethoxypyrimidinu

Průměrná hodnota odchylky u uhlíkového spektra je 1,84 ppm, což je mírně vyšší hodnota než u 1-methyluracilu. V případě vodíkového spektra je však průměrná odchylka menší než u 1-methyluracilu. Naměřené chemické posuny a vypočtené stínící konstanty vykazují velmi dobrou korelaci. Porovnání vypočtených a experimentálních hodnot NMR parametrů 2-amino-4,6-dimethoxypyrimidinu potvrdilo spolehlivost výpočetní metody.

4.4 Testování metody

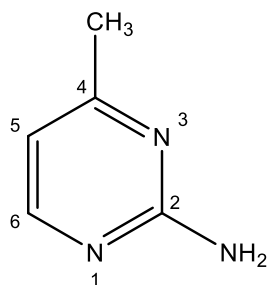
4.4.1 Detekce místa protonace 2-amino-4-methylpyrimidinu porovnáním experimentu s výpočtem

Metoda určení místa protonace byla otestována na molekule 2-amino-4-methylpyrimidinu (Obr. 22, str. 49), u kterého může docházet k protonaci na obou dusících pyrimidinového kruhu N1 a N3. Roztok této látky byl titrován z neutrálního do kyselého pH pomocí DCl. Po každém posunu hodnoty pH o přibližně 1 byla měřena ^1H NMR a ^{13}C NMR spektra.

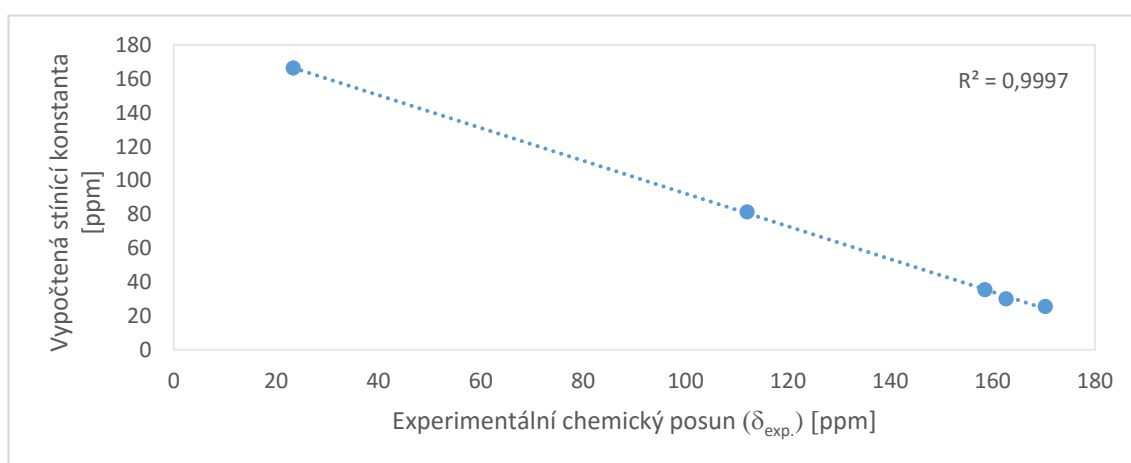
Optimalizace struktury a výpočet NMR parametrů byly provedeny u neutrální molekuly, N1 protonované formy i N3 protonované formy. Ve všech případech byla pro výpočty použita báze funkcí [6-31g(2df,2pd)] a simulace solvatace H_2O , při optimalizaci struktury i při výpočtu NMR parametrů.

Prvním krokem bylo přiřazení NMR signálů k jednotlivým atomům. V tomto případě byly všechny atomy C a H detekované a bylo možné je spolehlivě přiřadit pomocí spekter.

Korelace výpočet – experiment byla však taktéž analyzována pro kontrolu spolehlivosti (Obr. 23).

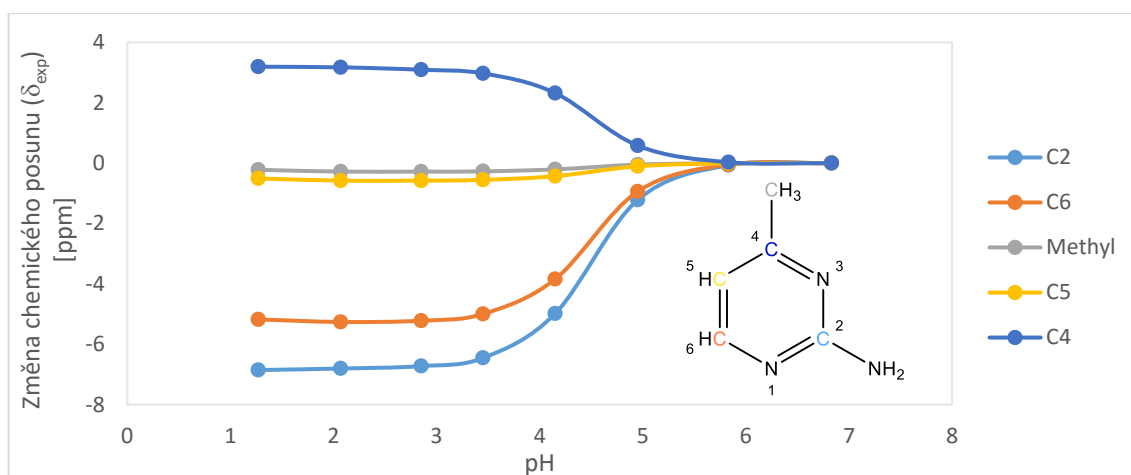


Obr. 22: 2-amino-4-methylpyrimidin



Obr. 23: Korelace výpočet – experiment pro NMR parametry ^{13}C 2-amino-4-methylpyrimidinu

Hodnoty vypočtených stínících konstant a experimentálních chemických posunů i v tomto případě velmi dobře korelují. Po přiřazení NMR signálů byly analyzovány změny jejich chemických posunů na titračních křivkách (Obr. 24, str. 50).



Obr. 24: Relativní chemické posuny uhlíků 2-amino-4-methylpyrimidinu v závislosti na pH

Stejně jako v případě 2-amino-4,6-dimethoxypyrimidinu se jedná o protonaci dusíku v aromatickém kruhu, takže by mělo dojít k výraznější záporné změně chemických posunů uhlíků sousedících s protonovaným dusíkem. K tomuto jevu skutečně došlo jak u uhlíku C2, tak u uhlíku C6. Můžeme tedy předpokládat, že k protonaci dochází právě na N1.

Dalším krokem bylo přiřazení experimentálních a vypočtených změn hodnot chemických posunů mezi neutrální a protonovanou formou (Tab. 4, str. 51).

Tab. 4: Porovnání experimentálních a vypočítaných hodnot změn chemických posunů mezi neutrální a ionizovanou formou 2-amino-4-methylpyrimidinu.

	$\delta_{\text{exp.}}$ pH 6,83	$\delta_{\text{exp.}}$ pH 1,27	$\Delta\delta_{\text{exp.}}$	$\Delta\delta_{\text{výpočet}}$ (N3)	Odchylka (N3)	$\Delta\delta_{\text{výpočet}}$ (N1)	Odchylka (N1)
¹³ C [ppm]							
C2	162,62	155,77	-6,85	-8,92	2,07	-10,25	3,40
C4	170,28	173,44	3,16	-7,25	10,41	13,05	9,89
C5	112,05	111,54	-0,51	0,79	1,30	0,77	1,28
C6	158,5	153,32	-5,18	9,14	14,32	-14,11	8,93
Methyl	23,35	23,13	-0,22	-3,89	3,67	2,52	2,74
Průměr					6,36		5,25
¹ H [ppm]							
H5	6,663	6,943	0,28	0,389	0,109	0,415	0,135
H6	8,115	8,292	0,177	0,536	0,359	-0,365	0,542
Methyl	2,315	2,538	0,223	0,333	0,110	0,420	0,197
Průměr					0,192		0,291

$\delta_{\text{exp.}}$: experimentálně získaný chemický posun; $\Delta\delta_{\text{exp.}}$: experimentálně získaný rozdíl chemických posunů mezi neutrální a protonovanou formou molekuly; $\Delta\delta_{\text{výpočet}} (N3)$: rozdíl chemických posunů mezi neutrální a N3 protonovanou formou molekuly získaný pomocí kvantově chemických výpočtů; odchylka (N3): absolutní hodnota rozdílu hodnot $\Delta\delta_{\text{výpočet}} (N3)$ a $\Delta\delta_{\text{exp.}}$; $\Delta\delta_{\text{výpočet}} (N1)$: rozdíl chemických posunů mezi neutrální a N1 protonovanou formou molekuly získaný pomocí kvantově chemických výpočtů; odchylka (N1): absolutní hodnota rozdílu hodnot $\Delta\delta_{\text{výpočet}} (N1)$ a $\Delta\delta_{\text{exp.}}$.

Aritmetické průměry odchylek vykazují příliš velké hodnoty oproti prvním dvěma modelovým molekulám. Mohlo by se jednat o chybu výpočtu, což však není v souladu s výsledkem korelace posunu neutrálních uhlíků, kde je chyba velmi malá. Druhým možným vysvětlením je přítomnost N3 protonované formy, která existuje při daném pH s N1 formou v dynamické rovnováze. Proto bylo potřeba najít poměr, ve kterém tyto dvě formy existují. Pro vypočítané hodnoty posunů byly nalezeny koeficienty, které by mohly určovat poměr jednotlivých forem tak, aby odchylka experimentu a váženého průměru výpočtů vynásobených těmito koeficienty, byla co nejnižší. Pro porovnání byl poměr obou forem v rovnováze vypočten i pomocí Boltzmanova zákona z volných Gibbsových energií jednotlivých forem, vypočítaných programem Gaussian. I při tomto poměru byly vypočteny průměrné odchylky změn chemických posunů experimentu a výpočtu. Porovnání výsledků je uvedeno v tabulce 5 na straně 52.

Tab. 5: Porovnání odchylky výpočet – experiment pro různá rovnovážná zastoupení N1 a N3 protonované formy.

	Minimalizace odchylky	Boltzmanův zákon
N1 prot. forma [%]	59	65
N3 prot. forma [%]	41	35
Průměrná odchylka ^{13}C [ppm]	1,28	1,66
Průměrná odchylka ^1H [ppm]	0,153	0,173

Minimalizace chyby: hodnoty získané pomocí minimalizace průměrné odchylky kvantově chemického výpočtu od experimentální hodnoty; Boltzmanův zákon: hodnoty získané dosažením vypočítaných volných Gibbsových energií do Boltzmanova zákona.

Hodnoty rovnovážného procentuálního zastoupení, při kterém by byla hodnota odchylky minimální, se velmi dobře shodují s hodnotami vypočítanými z Boltzmanova zákona. Potvrdila se tak spolehlivost výpočetní metody i v případě výpočtu rovnovážného zastoupení jednotlivých protonovaných forem.

Při rovnovážném zastoupení s minimálními odchylkami výpočtu od experimentu jsou hodnoty těchto odchylek velmi blízké modelovým příkladům. Porovnáním výsledků zvolené výpočetní metody s výsledky NMR experimentu se tedy s poměrně malou chybou podařilo určit primární místo protonace 2-amino-4-methylpyrimidinu, kterým je N1. Přítomnost N3 protonované formy však není vhodné zanedbat.

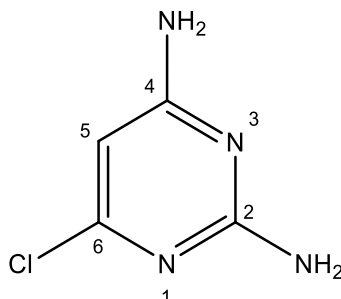
4.4.2 Detekce místa protonace báze s omezeným množstvím experimentálně dostupných dat.

Po optimalizaci a otestování byla metoda použita pro analýzu místa protonace molekuly 6-chloro-1,4-diaminopyrimidinu (Obr. 25, str. 53). Jedná se opět o molekulu, která může být protonována na N1 nebo N3 pyrimidinového kruhu, místo protonace není tedy jasně definováno strukturou.

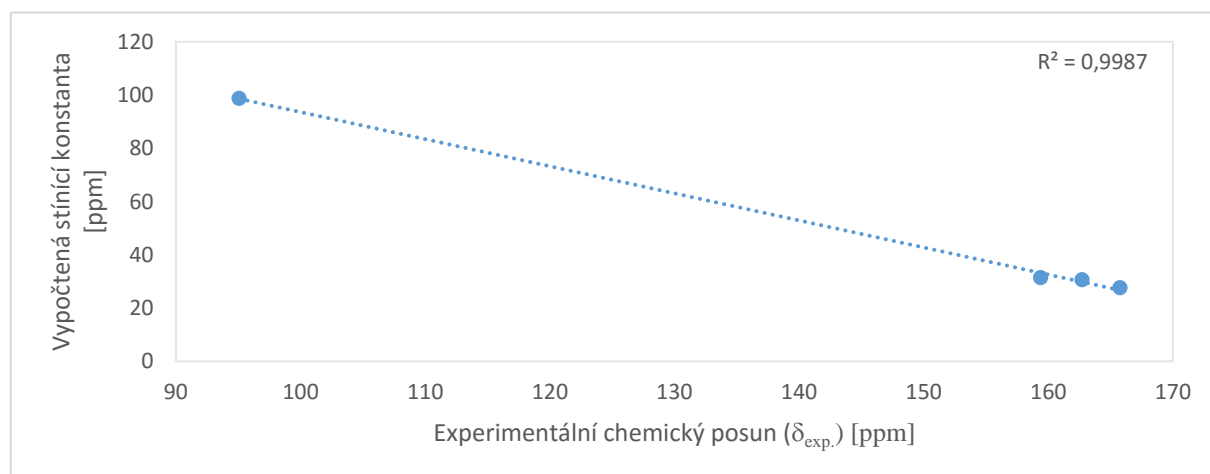
Optimalizace struktury a výpočet NMR parametrů byly provedeny u neutrální molekuly, N1 protonované formy i N3 protonované formy. Ve všech případech byla pro výpočet použita báze [6-31g(2df,2pd)] a simulace solvatace H_2O .

V tomto případě byly připraveny 2 roztoky, jeden neutrální a druhý silně kyselý, u kterých byla měřena ^{13}C NMR, ^1H NMR a $^1\text{H},^{13}\text{C}$ HMBC NMR spektra. Báze nebyla kompletně titrována, protože účelem bylo určit místo protonace s použitím minimálního množství experimentálních dat.

Jednoznačné přiřazení ^{13}C NMR signálů u této molekuly nebylo možné z důvodu interakce H5 se všemi kvartérními uhlíky v ^1H , ^{13}C HMBC spektru. Proto pro přiřazení signálu byla použita korelace výpočet-experiment (Obr. 26).



Obr. 25: 6-chloro-2,4-diaminopyrimidin



Obr. 26: Korelace výpočet – experiment pro NMR parametry ^{13}C 6-chloro-2,4-diaminopyrimidinu

Hodnota druhé mocniny vícenásobného korelačního koeficientu je velmi blízká jedné, což nasvědčuje tomu, že přiřazení pomocí korelace by mělo být spolehlivé. Jelikož jsou však signály kvartérních uhlíků od sebe vzdálené jen přibližně 3 ppm, existuje možnost, že došlo k jejich záměně a nesprávnému přiřazení. Na to, aby byly tyto chyby odhaleny, by bylo nutné provést podrobnější analýzu nebo získat více experimentálních dat.

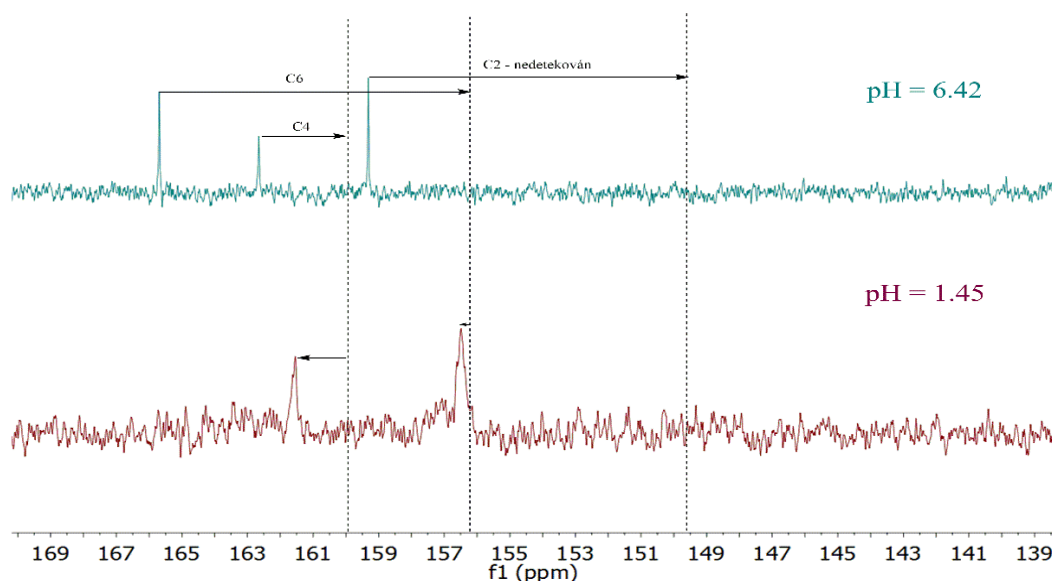
Chemické posuny atomů 6-chloro-2,4-diaminopyrimidinu (pH = 6,42): ^{13}C NMR (126 MHz, D_2O): $\delta = 165,76$ (C6); $162,72$ (C4); $159,39$ (C2); $95,07$ (C5) ppm. ^1H NMR (500 MHz, D_2O): $\delta = 6,030$ (s, 1H, H5) ppm.

Dalším problémem, který nastal, byla nepřítomnost jednoho signálu kvartérního uhlíku ve spektrech ^{13}C i ^1H , ^{13}C HMBC kyselého roztoku. Protože interakce nebyla

pozorována ani v ^1H , ^{13}C HMBC spektru, lze předpokládat, že jde o C2, uhlík nejvzdálenější od H5. Tento předpoklad není možné bez dalších informací nijak potvrdit. Z důvodu chybějícího signálu nebylo možné jednoznačně určit změny chemických posunů jednotlivých uhlíků pro porovnání s výpočtem. Proto byl zvolen postup vyzkoušený na 2-amino-4-methylpyrimidinu. Pomocí Boltzmanova zákona bylo z volných Gibbsových energií N1 a N3 protonované formy vypočteno jejich rovnovážné procentuální zastoupení.

Výsledné rovnovážné zastoupení: 78% N1 protonované a 22% N3 protonované formy.

Následně byl vypočten vážený průměr chemických posunů jednotlivých atomů těchto dvou forem. Každému atomu byla v průměru přidělena váha podle procentuálního zastoupení dané formy. Tento vážený průměr by měl, podobně jako u 2-amino-4-methylpyrimidinu, dobře odpovídat experimentálním chemickým posunům. Podle takto vypočítaných hodnot byly přiřazeny jednotlivé signály v ^{13}C spektru kyselého roztoku, a to přiřazením vypočtené hodnoty k nejbližší experimentální hodnotě (Obr. 27)



Obr. 27: Chemické posuny kvartérních uhlíků 6-chloro-2,4-diaminopyrimidinu v neutrálním a kyselém pH.

Jednotlivé šipky v neutrálním spektru znázorňují vypočítané změny chemických posunů. Pomocí těchto vypočítaných hodnot byly ve spektru měřeném při pH 1,45 přiřazeny signály.

Tato analýza souhlasí s předpokladem, že uhlík, který nebyl při kyselém pH detekovaný, je právě C2. Správnost přiřazení není však možné bez dalších měření potvrdit.

Pomocí takto přiřazených signálů byly opět vypočteny odchylky výpočet – experiment pro změnu chemického posunu mezi neutrální a protonovanou formou (Tab. 6, str. 55).

Tab. 6: Porovnání experimentálních a vypočítaných hodnot změn chemických posunů mezi neutrální a ionizovanou formou 6-chloro-2,4-diaminopyrimidinu.

Atom	$\delta_{\text{exp.}}$ pH 6,42	$\delta_{\text{exp.}}$ pH 1,45	$\Delta\delta_{\text{exp.}}$	$\Delta\delta_{\text{výpočet}}$	Odchylka
^{13}C [ppm]					
6C	165,76	156,50	-9,26	-9,34	0,08
4C	162,72	161,55	-1,17	-2,79	1,62
2C	159,39	-	-	-9,76	-
5C	95,07	95,84	0,77	4,33	3,56
Průměr	-	-	-	-	1,75
^1H [ppm]					
5H	6,030	6,234	0,204	0,431	0,227
Průměr	-	-	-	-	0,227

$\delta_{\text{exp.}}$: experimentálně získaný chemický posun; $\Delta\delta_{\text{exp.}}$: experimentálně získaný rozdíl chemických posunů mezi neutrální a protonovanou formou molekuly; $\Delta\delta_{\text{výpočet}}$: rozdíl chemických posunů mezi neutrální a protonovanou formou molekuly získaný pomocí kvantově chemických výpočtů; odchylka: absolutní hodnota rozdílu hodnot $\Delta\delta_{\text{výpočet}}$ a $\Delta\delta_{\text{exp.}}$.

Průměrné odchylky výpočtu od experimentu se přibližně shodují s odchylkami u modelových molekul.

Na základě malé odchylky můžeme předpokládat, že poměr forem vypočtených pomocí Boltzmanova zákona je blízký reálnému rovnovážnému poměru. K protonaci tedy dochází primárně na N1, přítomnost N3 protonované formy je ale opět nezanedbatelná.

Na základě těchto výsledků je možné konstatovat, že vypracovaná metodika má potenciál pro využití v praxi, a to i v případě experimentů s problematickými látkami. Zkoumané pyrimidinové báze představují skupinu látek s vlastnostmi i strukturou velmi blízkou bázím nukleových kyselin, a proto by metodika mohla být použita v řadě biochemických studií, ve kterých je detekce místa protonace bází důležitou informací. NMR parametry jednotlivých bází, získané pomocí kvantově-chemických výpočtů, vykazují jen malou chybu v porovnání s experimentálními hodnotami. Proto můžeme předpokládat, že chyba optimalizace geometrie a výpočtu volných Gibbsových energií je taktéž velmi malá a primárně protonované/deprotonované formy můžeme předpokládat výhradně na základě výpočtu.

Kombinací výpočtu s NMR spektroskopií, která je velmi účinnou a nedestruktivní analytickou metodou pro charakterizaci organických molekul, by mohly být jednoduše a rychle získány poměrně spolehlivé informace o acidobazických vlastnostech bází

nukleových kyselin. Tyto informace by mohly přispět k lepšímu pochopení mechanismů, kterých se účastní jejich ionizované formy. Mezi tyto biochemické procesy patří například acidobazická katalýza ribozymů. Při zkoumání těchto procesů byly dříve použity techniky založené na substituci jednotlivých bází a následném pozorování zvýšení/snížení rychlosti enzymové reakce.[104] Metodika vyvinutá v této práci by mohla pomoci s detekcí protonace/deprotonace přímo na bázích, které se účastní jednotlivých procesů a substituce by nebyla nutná. Detekce konkrétního místa protonace na pyrimidinovém kruhu představuje problém i při zkoumání vazby ligandu OxyTPP na TPP RNA přepínač.[76] Jak již bylo řečeno, místo protonace/deprotonace by taktéž bylo možné predikovat na základě struktury. Tato aplikace má potenciál například pro přípravu nových TFO nebo interkalátů DNA, jako je Triostin A, u kterého je posun pK_a cytosinů, a tedy jejich protonace, nevyhnutelnou podmínkou pro funkci.[99, 100]

Vypracovaná metoda má ale svá úskalí. Metoda byla testována na malém počtu velmi podobných molekul. U složitějších struktur by vzniklá chyba mohla být výrazně větší. Výsledné hodnoty procentuálního zastoupení jednotlivých ionizovaných forem navíc nebyly nijak potvrzeny, jedná se jen o přibližnou hodnotu. Pro měření vyžadující vysokou přesnost není tato metoda vhodná. U poslední studované molekuly se jedná hlavně o výsledky kvantově-chemických výpočtů, které nebylo možné experimentálně potvrdit. Jediným údajem pro kontrolu je hodnota chyby korelace výpočet – experiment, která souhlasí s modelovými látkami. Z těchto důvodů je nutné poznamenat, že pro zlepšení spolehlivosti a přesnosti je potřeba dalších experimentů.

5 Shrnutí

Protonace/deprotonace bází nukleových kyselin hraje důležitou roli v řadě biologických procesů. Sekundární struktury DNA, stabilizované neobvyklým párováním ionizovaných bází, jsou schopné modulovat expresi určitých genů či replikaci DNA. V některých případech jsou však takto stabilizované struktury zodpovědné za chyby v replikaci vedoucí ke genetickým poruchám. Protonace/deprotonace bází a existence minoritních tautomerních forem v oblasti aptameru RNA přepínačů zase ovlivňuje afinitu vazebního místa k určitým ligandům, takže se opět jedná o modulaci genové exprese. Ionizované báze některých RNA enzymů se zase přímo podílejí na acidobazické katalýze autolytických reakcí fosfodiesterové vazby, čímž simulují funkci ribonukleas. Ve všech případech je však podrobné studium mechanismů obtížné, jelikož protonace a deprotonace jsou velmi rychlé a dynamické procesy, které je náročné experimentálně detekovat.

Proto byla v této práci, pomocí korelace kvantově-chemických výpočtů a NMR spektroskopie, vyvinuta metodika detekce místa protonace/deprotonace na různých analogích pyrimidinových bází. Pomocí dvou modelů bází 1-methyluracilu a 2-amio-4,6-dimethoxypyrimidinu, u kterých je místo protonace/deprotonace jasně definováno strukturou, byla optimalizována výpočetní metoda tak, aby korelace experiment-výpočet vykazovala minimální odchylku NMR parametrů pro dané látky. Následně byla tato metoda otestována na bázích 2-amino-4-methylpyrimidinu a 6-chloro-2,4-diaminopyrimidinu, které obsahují ve svém aromatickém pyrimidinovém kruhu 2 dusíky. Ty se můžou protonovat, a proto primární místo protonace není definované strukturou. V obou případech bází s nejasným místem protonace bylo porovnáním vypočtených a naměřených hodnot určeno primární místo protonace a taktéž procentuální rovnovážné zastoupení jednotlivých protonovaných forem. Odchylka výpočtu a experimentu, velmi blízká modelovým příkladům, poukazuje na spolehlivost vypracované metodiky.

Jelikož byla metoda testovaná na molekulách strukturně velmi podobným bázím nukleových kyselin, má potenciál pro využití i ve studiu biochemických procesů, kterých se účastní ionizované formy. Lepší objasnění těchto procesů může mít velký přínos například v oblasti vývoje efektivních genových terapií nebo v biotechnologiích.

6 Seznam použité literatury

- [1] R. B. Corey, L. Pauling, *Arch. Biochem. Biophys.* 65, str. 164-81, 1956.
- [2] J. D. Watson, F. H. C. Crick, *Nature* 171, str. 737-738, 1953.
- [3] R. E. Franklin, R. G. Gosling, *Nature* 172, str. 156-157, 1953.
- [4] E. N. Nikolova, H. Zhou, F. L. Gottardo, H. S. Alvey, I. J. Kimsey, H. M. Al-Hashimi, *Biopolymers* 99, str. 955-968, 2013.
- [5] S. M. Mirkin, M. D. Frank-Kamenetskii, *Annu. Rev. Biophys. Biomol. Struct.* 23, str. 541-576, 1994.
- [6] A. C. Roth, *Bioinformatics* 28, str. I340-I348, 2012.
- [7] F. H. C. Crick, *J. Mol. Biol.* 19, str. 548-555, 1966.
- [8] G. Varani a W. H. McClain, *Embo Rep.* 1, str. 18-23, 2000.
- [9] V. J. Bardwell, M. Wickens, S. Bienroth, W. Keller, B. S. Sproat, A. I. Lamond, *Cell* 65, str. 125-133, 1991.
- [10] A. Matsuda, T. Sasaki, *Cancer Sci.* 95, str. 105-111, 2004.
- [11] A. Holy, H. Dvorakova, M. Masojdikova, *Collect. Czech. Chem. Commun.* 60, str. 1390-1409, 1995.
- [12] F. Eckstein, G. Gish, *Trends Biochem. Sci.* 14, str. 97-100, 1989.
- [13] R. Gan, X. Wu, W. He, Z. Liu, S. Wu, C. Chen, S. Chen, Q. Xiang, Z. Deng, D. Liang, S. Chen, L. Wang, *Sci. Rep.* 4, 2014.
- [14] P. F. Agris, " *Prog. Nucleic Acid Res. Mol. Biol.* 53, str. 79-129, 1996.
- [15] P. C. Bevilacqua, T. S. Brown, S. Nakano, R. Yajima, *Biopolymers* 73, str. 90-109, 2004.
- [16] B. Lippert, D. Gupta, *Dalton Trans.*, str. 4619-4634, 2009.
- [17] V. H. Harris, C. L. Smith, W. J. Cummins, A. L. Hamilton, H. Adams, M. Dickman, D. P. Hornby, D. M. Williams, *J. Mol. Biol.* 326, str. 1389-1401, 2003.
- [18] M. D. Topal, J. R. Fresco, *Nature* 263, str. 285-289, 1976.
- [19] D. Li, B. I. Fedeles, V. Singh, C. S. Peng, K. J. Silvestre, A. K. Simi, J. H. Simpson, A. Tokmakoff, J. M. Essigmann, *Proc. Natl. Acad. Sci. U. S. A.* 111, str. E3252-E3259, 2014.
- [20] W. N. Wang, H. W. Hellenga, L. S. Beese, *Proc. Natl. Acad. Sci. U. S. A.* 108, str. 17644-17648, 2011.

- [21] J. L. Wilcox, P. C. Bevilacqua, *J. Am. Chem. Soc.* 135, str. 7390-7393, 2013.
- [22] J. C. Cochrane, S. A. Strobel, *Acc. Chem. Res.* 41, str. 1027-1035, 2008.
- [23] A. Weixlbaumer, F. V. Murphy, A. Dziergowska, A. Malkiewicz, F. A. P. Vendeix, P. F. Agris, V. Ramakrishnan, *Nat. Struct. Mol. Biol.* 14, str. 498-502, 2007.
- [24] G. K. Mittapalli, K. R. Reddy, H. Xiong, O. Munoz, B. Han, F. De Riccardis, R. Krishnamurthy, A. Eschenmoser, *Angew. Chem.* 46, str. 2470-2477, 2007.
- [25] P. O. P. Tso, I. S. Melvin, A. C. Olson, *J. Am. Chem. Soc.* 85, str. 1289-1296, 1963.
- [26] P. O. P. Tso, *Ann. N. Y. Acad. Sci.* 153, str. 785-804, 1969.
- [27] B. Lippert, *Chem. Biodiversity* 5, str. 1455-1474, 2008.
- [28] J. L. Wilcox, A. K. Ahluwalia, P. C. Bevilacqua, *Acc. Chem. Res.* 44, str. 1270-1279, 2011.
- [29] K. Hoogsteen, *Acta Crystallogr.* 12, str. 822-823, 1959.
- [30] Y. Courtois, P. Fromageo, W. Guschlba, *Eur. J. Biochem.* 6, str. 493-501, 1968.
- [31] B. L. Haas, W. Guschlbauer, *Nucleic Acids Res.* 3, str. 205-218, 1976.
- [32] R. M. Izatt, J. Christen, J. H. Rytting, *Chem. Rev.* 71, str. 439-481, 1971.
- [33] J. L. Asensio, A. N. Lane, J. Dhesi, S. Bergqvist, T. Brown, *J. Mol. Biol.* 275, str. 811-822, 1998.
- [34] G. E. Plum, Y. W. Park, S. F. Singleton, P. B. Dervan, K. J. Breslauer, *Proc. Natl. Acad. Sci. U. S. A.* 87, str. 9436-9440, 1990.
- [35] M. D. Frank-Kamenetskii, S. M. Mirkin, *Annu. Rev. Biochem.* 64, str. 65-95, 1995.
- [36] Q. Wu, S. S. Gaddis, M. C. MacLeod, E. F. Walborg, H. D. Thames, J. DiGiovanni, K. M. Vasquez, *Mol. Carcinog.* 46, str. 15-23, 2007.
- [37] S. M. Mirkin, V. I. Lyamichev, K. N. Drushlyak, V. N. Dobrynin, S. A. Filippov, M. D. Frankkamenetskii, *Nature* 330, str. 495-497, 1987.
- [38] V. I. Lyamichev, S. M. Mirkin, M. D. Frank-Kamemetskii, *J. Biomol. Struct. Dyn.* 3, str. 667-669, 1986.
- [39] J. C. Hanvey, E. M. Williams, J. M. Besterman, *Antisense Res. Dev.* 1, str. 307-17, 1991.
- [40] B. S. Rao, *Mol. Cell. Biochem.* 156, str. 163-168, 1996.
- [41] R. D. Wells, D. A. Collier, J. C. Hanvey, M. Shimizu, F. Wohlrab, *FASEB J.* 2, str. 2939-2949, 1988.

- [42] S. Nakano, D. M. Chadalavada, P. C. Bevilacqua, *Science* 287, str. 1493-1497, 2000.
- [43] Y. I. Kuzmin, C. P. Da Costa, J. W. Cottrell, M. J. Fedor, *J. Mol. Biol.* 349, str. 989-1010, 2005.
- [44] P. C. Durant, D. R. Davis, *J. Mol. Biol.* 285, str. 115-131, 1999.
- [45] G. Chen, S. D. Kennedy, D. H. Turner, *Biochemistry* 48, str. 5738-5752, 2009.
- [46] P. Fojtik, M. Vorlickova, *Nucleic Acids Res.* 29, str. 4684-4690, 2001.
- [47] M. Vorlickova, I. Kejnovska, M. Tumova, J. Kypr, *Eur. Biophys. J.* 30, str. 179-185, 2001.
- [48] G. J. Quigley, G. Ughetto, G. A. Vandermarel, J. H. Vanboom, A. H. J. Wang, A. Rich, *Science* 232, str. 1255-1258, 1986.
- [49] J.-W. Keum, H. Bermudez, *Chem. Commun.* 48, str. 12118-12120, 2012.
- [50] A. K. Jissy, A. Datta, *J. Phys. Chem. B* 114, str. 15311-15318, 2010.
- [51] M. J. Behe, *Nucleic Acids Res.* 23, str. 689-695, 1995.
- [52] R. Zain, J. S. Sun, *Cell. Mol. Life. Sci.* 60, str. 862-870, 2003.
- [53] M. Kato, N. Shimizu, *J. Biochem.* 112, str. 492-494, 1992.
- [54] G. Duvalvalentin, T. Debizemont, M. Takasugi, J. L. Mergny, E. Bisagni, C. Helene, *J. Mol. Biol.* 247, str. 847-858, 1995.
- [55] Y. Kohwi, T. Kohwishigematsu, *Genes Dev.* 5, str. 2547-2554, 1991.
- [56] A. K. Maiti, S. K. Brahmachari, *BMC Mol. Biol.* 2, str. 11-11, 2001.
- [57] G. Xu, A. G. Goodridge, *Arch. Biochem. Biophys.* 358, str. 83-91, 1998.
- [58] M. J. Ulrich, W. J. Gray, T. J. Ley, *J. Biol. Chem.* 267, str. 18649-18658, 1992.
- [59] G. J. Mavrothalassitis, D. K. Watson, T. S. Papas, *Oncogene* 5, str. 1337-1342, 1990.
- [60] S. Q. Sun, E. S. Stoflet, J. G. Cogan, A. R. Strauch, M. J. Getz, *Mol. Cell. Biol.* 15, str. 2429-2436, 1995.
- [61] D. G. Pestov, A. Dayn, E. Y. Siyanova, D. L. George, S. M. Mirkin, *Nucleic Acids Res.* 19, str. 6527-6532, 1991.
- [62] M. Santra, K. G. Danielson, R. V. Iozzo, *J. Biol. Chem.* 269, str. 579-587, 1994.
- [63] E. K. Hoffman, S. P. Trusko, M. Murphy, D. L. George, *Proc. Natl. Acad. Sci. U. S. A.* 87, str. 2705-2709, 1990.

- [64] K. Zahedi, J. J. Bissler, A. E. Prada, J. A. Prada, A. E. Davis, *J. Immunol.* 162, str. 7249-7255, 1999.
- [65] G. M. Samadashwily, A. Dayn, S. M. Mirkin, *Embo J.* 12, str. 4975-4983, 1993.
- [66] V. Lyamichev, M. A. D. Brow, J. E. Dahlberg, *Science* 260, str. 778-783, 1993.
- [67] M. Peleg, V. Kopel, J. A. Borowiec, H. Manor, *Nucleic Acids Res.* 23, str. 1292-1299, 1995.
- [68] B. S. Rao, H. Manor, R. G. Martin, *Nucleic Acids Res.* 16, str. 8077-8094, 1988.
- [69] J. F. Lee, J. R. Hesselberth, L. A. Meyers, A. D. Ellington, *Nucleic Acids Res.* 32, str. D95-D100, 2004.
- [70] W. Winkler, A. Nahvi, R. R. Breaker, *Nature* 419, str. 952-956, 2002.
- [71] B. J. Tucker, R. R. Breaker, *Curr. Opin. Struct. Biol.* 15, str. 342-348, 2005.
- [72] M. Mandal, R. R. Breaker, *Nat. Struct. Mol. Biol.* 11, str. 29-35, 2004.
- [73] S. D. Gilbert, F. E. Reyes, A. L. Edwards, R. T. Batey, *Structure* 17, str. 857-868, 2009.
- [74] S. Thore, M. Leibundgut, N. N. Ban, *Science* 312, str. 1208-1211, 2006.
- [75] S. Thore, C. Frick, N. Ban, *J. Am. Chem. Soc.* 130, str. 8116-8117, 2008.
- [76] V. Singh, C. S. Peng, D. Li, K. Mitra, K. J. Silvestre, A. Tokmakoff, J. M. Essigmann, *ACS Chem. Biol.* 9, str. 227-236, 2014.
- [77] P. C. Bevilacqua, T. S. Brown, D. Chadalavada, J. Lecomte, E. Moody, S. Nakano, *Biochem. Soc. Trans.* 33, str. 466-470, 2005.
- [78] L. Simon J. M. Goodman, *J. Org. Chem.* 75, str. 1831-1840, 2010.
- [79] M. J. Fedor, *J. Mol. Biol.* 297, str. 269-291, 2000.
- [80] F. D. Jones S. A. Strobel, *Biochemistry* 42, str. 4265-4276, 2003.
- [81] M. D. Been G. S. Wickham, *Eur. J. Biochem.* 247, str. 741-753, 1997.
- [82] W. C. Winkler, A. Nahvi, A. Roth, J. A. Collins, R. R. Breaker, *Nature* 428, str. 281-286, 2004.
- [83] R. T. Raines, *Chem. Rev.* 98, str. 1045-1065, 1998.
- [84] P. B. Rupert, A. R. Ferre-D'Amare, *Nature* 410, str. 780-786, 2001.
- [85] M. Martick, W. G. Scott, *Cell* 126, str. 309-320, 2006.

- [86] N. B. Suslov, H. Huang, D. M. J. Lilley, P. Rice, J. A. Piccirilli, *Abstr. Pap. Am. Chem. Soc.* 245, 2013.
- [87] I. Velikyan, S. Acharya, A. Trifonova, A. Foldesi, J. Chattopadhyaya, *J. Am. Chem. Soc.* 123, str. 2893-2894, 2001.
- [88] S. R. Das, J. A. Piccirilli, *Nat. Chem. Biol.* 1, str. 45-52, 2005.
- [89] J. Han, J. M. Burke, *Biochemistry* 44, str. 7864-7870, 2005.
- [90] J. C. Cochrane, S. V. Lipchick, S. A. Strobel, *Chemistry & Biology* 14, str. 97-105, 2007.
- [91] L. L. Lebruska, Kuzmine, M. J. Fedor, *Chem. Biol.* 9, str. 465-473, 2002.
- [92] P. Nissen, J. Hansen, N. Ban, P. B. Moore, T. A. Steitz, *Science* 289, str. 920-930, 2000.
- [93] G. M. Samadashwily, G. Raca, S. M. Mirkin, *Nat. Genet.* 17, str. 298-304, 1997.
- [94] C. T. McMurray, *Chromosoma* 104, str. 2-13, 1995.
- [95] K. Ohshima, R. D. Wells, *J. Biol. Chem.* 272, str. 16798-16806, 1997.
- [96] K. Ohshima, L. Montermini, R. D. Wells, M. Pandolfo, *J. Biol. Chem.* 273, str. 14588-14595, 1998.
- [97] R. R. Sinden, *Am. J. Hum. Genet.* 64, str. 346-353, 1999.
- [98] J. S. Sun, C. Helene, *Curr. Opin. Struct. Biol.* 3, str. 345-356, 1993.
- [99] D. Praseuth, A. L. Guieysse, C. Helene, *Biochim. Biophys. Acta, Gene Struct. Expression* 1489, str. 181-206, 1999.
- [100] A. H. J. Wang, G. Ughetto, G. J. Quigley, A. Rich, *J. Biomol. Struct. Dyn.* 4, str. 319-342, 1986.
- [101] L. Song, V. H. B. Ho, C. Chen, Z. Yang, D. Liu, R. Chen, D. Zhou, *Adv. Healthcare Mater.* 2, str. 275-280, 2013.
- [102] Y. Dong, Z. Yang, D. Liu, *Acc. Chem. Res.* 47, str. 1853-1860, 2014.
- [103] G. C. K. Roberts, *NMR of macromolecules : a practical approach*. Oxford University Press, 1993.
- [104] S. Kath-Schorr, T. J. Wilson, N.-S. Li, J. Lu, J. A. Piccirilli, D. M. J. Lilley, *J. Am. Chem. Soc.* 134, str. 16717-16724, 2012.

„Svoluji k zapůjčení této práce pro studijní účely a prosím, aby byla řádně vedena evidence
vypůjčovateli.“

Jméno a příjmení, adresa	Číslo OP	Datum vypůjčení	Poznámka